

Time to take the plunge at Yucca Mountain

The process from which the Nevada desert has emerged as the only candidate site for disposal of US high-level nuclear waste has been flawed. But the result may still be a safe repository.

Yucca Mountain in Nevada is, from a certain perspective, a splendid site for the disposal of the 100,000 tonnes of high-level nuclear waste that the United States civil and military nuclear programmes have already produced or will produce early in the next century. The Nevada mountain is a dry slab of desert rock, where the waste can sit far beneath the surface and far above the water table. By the standards of Europe or of the eastern United States, it is extraordinarily isolated. There weren't many people around to complain when the United States government set off several hundred nuclear explosions at the adjoining Nevada Test Site.

The US Department of Energy (DoE) is currently engaged in an extremely expensive and thorough assessment of the site's suitability for the storage of steel cylinders containing ceramic spent nuclear fuel, as well as vitrified waste from the nuclear weapons programme, and is due to decide in 2001 whether to apply for a licence to construct a repository at Yucca.

This decision is of critical importance both to the nuclear power industry, which has nowhere else to put its waste, and to the DoE, which badly needs to show its critics that it is capable of exercising its functions properly. It will play an equally critical role in demonstrating whether the US government is able to convince its own people that it can be trusted to handle such matters responsibly.

President Bill Clinton has so far handled the matter with the mixture of political agility and practical inactivity which so characterizes his administration. By holding out for the 2001 decision, and repeatedly vetoing legislation that would immediately build an interim storage site at Yucca Mountain, he has managed to champion due process, to avoid upsetting his allies in the environmental movement, and to defer any meaningful decision to his successor, all at once.

Mounting pressure

But like all deferred choices, this one has costs. With every year that passes assessing a site to which the DoE can offer no alternative, the political pressure for an outcome increases. The DoE has, since last year, been legally obliged to accept spent fuel from power generators who have been paying a special levy towards the cost of a repository. A barrage of legal actions from the generators and their state governments are now under way in response to the resultant breach of contract. The DoE's plans for clean-up of the disused nuclear weapons complex will also be thrown into disarray by any decision not to proceed with the repository.

Yet the department has no site options under consideration, except Yucca. Furthermore, Yucca Mountain was never selected on technical grounds in the first place. Rather, it was unfortunate enough to lose a game of musical chairs a decade ago, when all of the other site options managed to find grounds (sometimes technical, but usually political) to have themselves ruled out of the running. It is no wonder that opponents of the Yucca Mountain plan refuse to believe that the DoE is ready to judge it fairly on technical grounds.

A report released last week (see page 500) suggests that the moun-

tain cannot be relied upon to remain dry in the very long term. Further alleged technical problems with the site include its geological complexity, the possibility that conditions there will attack and erode the glass waste after the steel surrounding it has eroded away, and that radioactive carbon-14 in the waste will oxidize and escape into the atmosphere.

But rather than developing a technical case, which the state of Nevada certainly had the means and the access to do, opponents are likely to depend on public sentiment and political brawn to attempt to bring about Yucca's downfall. The state has just re-elected Senator Harry Reid, largely on the grounds that he, as a high-ranking Democrat in the US Senate, will be well positioned to obstruct the project.

Public concern

Reid and his allies have until now been able to muster the 34 Senate votes needed to prevent that body from overturning Clinton's veto on the interim storage proposal, despite an immense lobbying effort by the nuclear industry in favour of the proposal. That effort will be repeated in the new year, after the DoE publishes a 'viability assessment' of Yucca that will probably say that the department has identified no overwhelming technical problems at the site. However, the balance of power in the Senate is unchanged, and a Clinton veto of interim storage is unlikely to be overruled.

That will leave two years of debate before the decision in 2001. During that time, public concern about the proposal will doubtless increase. Las Vegas has expanded, since the days of the first nuclear tests, from a one-horse town to the fastest-growing city in America, and its powerful casino industry will fight tooth and nail to oppose a project that it thinks has the potential to slow the hitherto insatiable flow of tourists to its gambling palaces.

The process by which Yucca Mountain has been chosen as the only candidate repository site has been far from perfect. But the technical assessment of the site that the DoE is now undertaking is exceptionally open and thorough, whether measured by international standards or those previously used in the United States.

Opponents of the project are tarnished by their own failure to offer an alternative solution to the nuclear waste problem. Some of the environmental groups opposed to the project appear primarily interested in reaching a gridlock that achieves their real objective, namely the complete undermining of the feasibility of existing nuclear power plants. Permanent storage of spent fuel at Yucca Mountain is what the nuclear industry wants and believes that it is already paying for, through the levy. At this time, it appears more likely that a repository will indeed be built at the Nevada site, but that it will store spent fuel in a form that will allow it to be retrieved at a later date. The alternatives offered so far, such as the transmutation of waste under a bombardment of neutrons, still face major technical hurdles. If opponents of the Yucca Mountain project cannot explain what alternative approach they support they must not be surprised if the project gets the go-ahead in three years' time. □



Patent on gene fragment sends researchers a mixed message...

[WASHINGTON] The award of the first US patent for expressed sequence tags (ESTs) — short sequences of DNA commonly used as tools to decode longer sequences — has led to speculation that similar rights might soon be granted to gene fragments whose patents have been pending for years.

The patent was awarded in October to Incyte Pharmaceuticals, Inc. of Palo Alto, California. But some in the biotechnology community say the award may not be relevant to most EST patent applications. Indeed, there is some doubt as to whether the Patent and Trademark Office even considered Incyte's inventions to be ESTs.

The controversy over ESTs began in the early 1990s when Craig Venter, then a researcher with the National Institutes of Health (NIH), and subsequently the founder of the Institute for Genomic Research, pioneered their use as a shortcut to sequencing the human genome.

At first, the NIH applied for patents on ESTs. But after much political turmoil, it began to oppose proprietary rights to gene fragments, and raised its concern last year at the news that the patent office was likely to grant such patents (see *Nature* 386, 312; 1997). Meanwhile, however, private companies flooded the office with applications covering millions of ESTs.

The patent granted to Incyte is for one full-length gene, as well as polynucleotides



that encode more than 40 protein kinases. Lee Bendekgey, the company's vice president for legal affairs, says this was not the company's first application for an EST patent, but it "breezed through relatively easily".

The reason, he believes, was the full-length gene, and that the fragments were highly characterized. In early debates about EST patenting, Bendekgey says, they were stereotyped as being almost random DNA sequences about which little was known.

That may still be true for many of the 1.2 million ESTs for which Incyte has filed patent applications, but Bendekgey and others say it no longer holds for other fragments with better known functions.

Dave Schmickel, a patent counsel with the Biotechnology Industry Organization, says the term EST has become somewhat vague: "An EST is in the eye of the beholder". This is why, according to Patent Office spokesperson Brigid Quinn, some people would even debate whether Incyte's invention "was in fact an EST".

Bendekgey says the office expressed its displeasure when the company issued a press release in November touting the award as the first patent for an EST.

Schmickel and others say that more important than terminology in predicting the success of EST patent applications is whether a company tries to claim broader rights than are warranted by its knowledge of the fragment's structure and function.

Draft guidelines for biotechnology patents, published by the patent office in the summer, say there should be a "correspondence between what the applicant has described... and what the applicant is claiming".

That may already rule out what some biotechnology researchers fear — that the owner of an EST, even one whose function is unknown, would automatically gain patent rights to larger fragments or whole genes containing that EST that are discovered later.

Any patented invention must pass several tests; for example, it must have a known function, must be novel and non-obvious (meaning that not anyone can work it out based on the current state of the art), and be adequately described. Claiming rights to a gene of unknown function, based on a subset of the whole gene, would not appear to pass these tests.

Not all the fears have been eased, however. At a meeting on intellectual property rights at the National Academy of Sciences in Washington last week, Francis Collins, head of the NIH's National Human Genome Research Institute, called the Incyte patent award a "disturbing turn of events" and feared that it signalled the patent office's willingness to grant overly broad proprietary rights.

But Jack Tribble, a patent lawyer with Merck and Co., predicted at the same meeting that EST patents will not preclude a patent on the parent gene being granted to someone else, unless perhaps the gene contains a large number of proprietary ESTs, a matter that may have to be resolved in the courts.

He and others also predict that only a fraction of the millions of EST patent applications on file will be granted. Companies and universities will lack the time and money to pursue — and defend — patents on all but their top-priority gene fragments. **Tony Reichhardt**

... as Germany hesitates over Brussels directive

[WASHINGTON] Germany's new left-leaning coalition government might be a stumbling block to the approval of laws governing biotechnology patents, according to an official at the country's Ministry of Justice.

European states have until July 2000 to draw up laws implementing last May's European Union directive on biotechnological inventions (see *Nature* 393, 299; 1998). The directive allows patents on isolated human genes and gene fragments with known functions, as well as on transgenic life.

But Hans-Georg Landfermann, head of the ministry's subdivision on industrial property protection

and copyright law, said last week that "it will not be easy for Germany to comply with the [July 2000] deadline". The coalition government of Social Democrats and Greens elected in October will probably disapprove of any law that gives broad property rights over genetically engineered organisms. His remarks came at a symposium on international intellectual property rights at the National Academy of Sciences in Washington.

The EU directive was passed after years of controversy, during which the European Patent Office imposed a moratorium on the patenting of plants and

animals. Greens strongly opposed the new directive while it was being debated, Landfermann says, and will "scrutinize every word of the draft [German] law". The Dutch government has appealed to the European Court of Justice to annul the directive, although the appeal is expected to fail (see *Nature* 395, 736; 1998).

Landfermann said the justice ministry draft will probably be similar to the EU directive. But, he told the symposium, "don't expect much from the German draft law," which will be "politically difficult" to pass. Meanwhile, he said, the German patent office would follow the new EU policies provisionally. **T.R.**

Planned US nuclear repository 'at risk' ...

[WASHINGTON] The US Department of Energy (DoE) is preparing to investigate claims that warm underground fluids may have entered its proposed nuclear waste repository at Yucca Mountain, Nevada, at some time in the past, possibly less than a million years ago.

Scientists from the US Geological Survey (USGS) and the University of Las Vegas at Nevada are to collect, analyse and date calcite samples from the site. Calcite is a signature of the historical presence of water.

The claims are made in a study published last week by the independent Institute for Energy and Environmental Research based near Washington DC. The study's author is Yuri Dublyansky, a geologist with the Russian Academy of Sciences, and former consultant to the state of Nevada.

Dublyansky has been studying the Yucca Mountain site for the past five years. The state is opposed to a repository on its soil and, if confirmed, Dublyansky's claims could raise further questions about the suitability of Yucca Mountain as a repository for 77,000 tonnes of civilian and military nuclear waste.

This is because one of the site's main attractions when it was chosen 12 years ago was the belief, backed up by a strong scientific consensus, that the site has been dry for at least 8 million years.

Water is among the proposed repository's worst enemies, as it will corrode the metal containers used to store the waste. This will lead to radionuclides leaking out and eventually contaminating the environment.

USGS scientists are sceptical of Dublyansky's claims, and say they do not expect their own analysis to provide confirmatory evidence. But leading geologists from other universities in the United States

and Europe are more supportive.

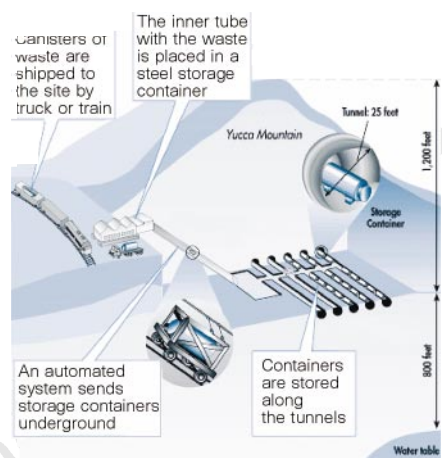
Robert Loux, executive director of the Nevada state Nuclear Waste Projects Office, says that if the age of the samples is found to be around 100,000 years old, the basis for choosing Yucca Mountain would be called into question.

This is also the view of Congress's scientific advisory board on nuclear waste, the Nuclear Waste Technical Review Board. In an assessment of some of Dublyansky's preliminary work, the board concluded: "If fluids at elevated temperatures were present less than 100,000 years ago, this could lend credence to the hypothesis of ongoing hydrothermal activity at Yucca Mountain."

This possible setback to the DoE's plans comes after criticisms made last month by the board. In its latest report to Congress, the board says there are still many uncertainties in the site's suitability as a repository for nuclear waste. The report also calls on the DoE to consider alternatives to its proposed designs for the repository (see below).

These developments could not have come at a worse time for the DoE, which is due to present its design ideas for the repository, and an estimate of costs, to Congress by the end of this month. The department has until 2001 to decide if it is to recommend the site to the president.

An application for a construction licence will need the approval of both the president and Congress. Dublyansky says his conclusion that warm fluids once ascended into the repository are based on an analysis of the gases and vapour found trapped in calcite samples taken from a five-mile tunnel that has been excavated at the site for research. The gases are known as 'fluid inclusions'.



The Yucca Mountain storage plan: its stability in the face of geological change is being questioned.

Dublyansky says that the temperature of the samples at the point of entrapment was between 35 and 75 °C. "Water at such a temperature could not have come from surface sources," he claims.

But Joe Whelan of the USGS says that Dublyansky's inference that the calcite must therefore have formed from warm, ascending underground fluids is pure speculation. USGS scientists believe the water probably percolated downwards from rain or snow.

Whelan and six colleagues accuse the Russian scientist of ignoring contrary conclusions from a 1992 study by the National Research Council, as well as the absence of any known physical mechanism by which the fluids might have risen.

"This is not a scholarly investigation," says Whelan of Dublyansky's work. "Strong but careless conclusions are drawn from off-hand observations, and possible interpretations of data are presented as established fact, with no effort to objectively consider plausible alternatives."

Arjun Makhijani, president of the Institute for Energy and Environmental Research, says he was taken aback by the nature and tone of the USGS response. But he says that this was the exception out of five reviews of Dublyansky's work.

Three of the other four reviewers, including Bruce Yardley, professor of earth sciences at the University of Leeds, UK, and Larry Diamond, professor of mineralogy and petrology at the University of Leoben, Austria, support Dublyansky's conclusion.

The fourth reviewer, Jean Cline of the University of Las Vegas at Nevada, is more cautious. She says the fluid-inclusion data by themselves suggest no more than that fluids at elevated temperatures once existed at Yucca Mountain. More research, she says, is needed to understand fully how they were formed.

Ehsan Masood

... as review board asks DoE to think again

[WASHINGTON] The US Department of Energy favours placing nuclear waste inside 8-metre-wide tunnels excavated in a dry part of Yucca Mountain (see above). The tunnels will be 300 metres below the land surface, and 300 metres above a water table. The waste will be encased in steel tubing coated with a two-centimetre-thick layer of a special, corrosion-resistant nickel alloy.

Congress's scientific advisory board on nuclear waste says alternatives to the nickel-alloy coating need to be investigated. The shell

is designed to resist the eventual corrosion from the outer steel shell by forming a thin layer between itself and the steel. But the board says that the layer itself may be vulnerable to other forms of damage.

William D. Barnard, the board's executive director, says the board also believes that the waste should be stored in such a way that temperatures can be maintained below 100 °C. Under the DoE's planned design, the waste will be densely stored, generating temperatures of around 200 °C.

These temperatures, the board says, may cause water above the tunnels – possibly having percolated from rain or snowfall – to boil. It would then vaporize, cool down, condense and possibly flow downwards in the direction of the tunnels.

Allen Benson, a spokesman for the department, says that an eight-year research programme is underway to assess the effect of high temperatures on the flow of water through the rock. But he says it is unlikely that the design of the metal tubing will change.

E.M.

Resignations hit French science reforms

[PARIS] Plans by Claude Allègre, France's science minister, for sweeping reforms of the country's research system suffered a serious setback last week when the entire senior administration of his ministry's research department resigned *en bloc*.

The ministry has put a brave face on the unprecedented event. Daniel Nahon, director general of the department, explained that the seven directors who have offered their resignations want to leave for a variety of reasons, including simply wishing to return to their laboratories after a year of hard work at the ministry. (Nahon himself is also reported to have resigned, but he could not be contacted last week for confirmation.)

Pierre-Louis Curien, director of physics and engineering sciences at the ministry, says that one reason he resigned is to be able to return to research. But he adds that, although the ministry has identified important goals, its operational structures, new central research funds, and the structures established to coordinate the research agen-

cies and universities, "have not yet converged to a setting that works satisfactorily" (see *Nature* 395, 422; 1998).

Curien says: "I hope very much that others will be able to carry out all this important work efficiently, and that, to begin with, a better picture of 'who does what' will be stabilized as soon as possible."

One explanation put forward for the mass resignation is confusion within the ministry over responsibilities. The resignations seem to have been partly prompted by the creation of a 'mission' for university research and doctoral studies within the department of research. This new entity is perceived by many in the department as duplicating much of the department's own brief, and so likely to provoke infighting.

The resignations are all the more embarrassing for the ministry because Allègre has made reform of his central administration the centrepiece of his overall reform strategy. The reorganization has streamlined the ministry from 19 departments to 11, with money

and power being concentrated in three: research, higher education and technology.

Jacques Fossay, a member of the main trade union representing scientists (SNCS), claims that the resignations mark a "disavowal" of Allègre's reform plans by those who would have had responsibility for implementing them. Fossay argues that it is difficult to believe that such drastic and symbolic action could be explained simply by dissatisfaction with organizational matters within the ministry.

Fossay says that those resigning cannot have failed to notice that the move comes at a critical time for Allègre. His reform plans face a revolt next week, when the National Committee for Scientific Research — an elected body of scientists that governs much of French research — will hold a plenary meeting to discuss the future of research, and to oppose Allègre's plans to reform the Centre National de la Recherche Scientifique, the country's basic research agency (see *Nature* 395, 729–730; 1998). **Declan Butler**

Japanese industry defies recession to spend more on research

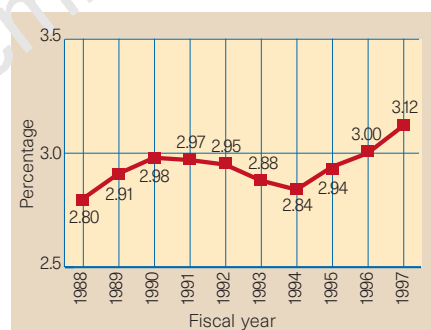
[TOKYO] Despite Japan's severe economic situation, spending on industrial research and development increased by a record amount in the fiscal year 1997, according to figures released last week by Tokyo's Management and Coordination Agency.

The agency's annual report on R&D expenditure reveals that Japan's overall research spending grew by 4.4 per cent over the 1997 fiscal year, which ended in March. Total spending on R&D reached 3.12 per cent of the nation's gross national product — considerably higher than that of all other large industrialized nations (the United States, for example, spent 2.43 per cent of its gross national product on research in 1996).

The increased expenditure, however, only occurred in 'developmental costs', which account for 61.7 per cent of the overall R&D expenditure. In contrast, spending on basic research, which took up 13.8 per cent of total expenditure and is carried out primarily in universities and research institutes, has decreased from last year.

Although industry spent a modest 6.2 per cent on basic research, the figure has grown by 7.4 per cent from the previous fiscal year. But spending on basic research at universities and research institutes fell slightly from last year.

There was increased support for R&D from all sectors, especially from industry, which faced considerable declines in both revenue and profits in 1997. Although such declines are continuing, many companies say their investment in R&D is unlikely to suffer.



New heights: Japan is increasing the proportion of gross national product spent on R&D.

The report is based on a survey of more than 15,000 companies, universities and research institutes across Japan. It estimates total R&D expenditure at 15.7 trillion yen (US\$130 billion). Spending on industrial R&D, excluding the software industry, grew by 6.1 per cent, while universities and research institutes showed modest increases of 0.8 and 1.5 per cent, respectively.

Overall private spending on research, which accounts for almost four-fifths of total R&D expenditure, increased by 4.9 per cent, with a marked increase in the electronics, telecommunications and transport sectors.

In contrast, however, public research spending grew by only 1.4 per cent, leading many to doubt whether Japan will reach its goal for the five-year plan for science and technology, launched in 1996, of doubling government spending on science by 2001.

The discrepancy between the figure displayed in the survey and the projected increase of 11.9 per cent in government research funding budgeted for the 1997 fiscal year (see *Nature* 385, 104; 1997) can be partly explained by generous supplementary budgets for the previous fiscal year.

Although supplementary budgets appeared to add considerably more to overall public research spending, the actual amount spent on research was minimal, and thus the target for the five-year plan is still likely to be out of reach.

A survey carried out by *Nihon Keizai Shinbun*, a financial daily newspaper, shows that major listed companies in Japan intend to spend even more on R&D in the current fiscal year. Sony plans to increase worldwide R&D spending by ¥350 billion, while the Matsushita Electric Industrial Company projects a 4.1 per cent rise to ¥500 billion.

"Although profit and revenue declines are expected to continue into the current fiscal year, spending on R&D will be given priority, as we consider research to be the lifeline of the company," says Akira Kadota, a spokesman for Matsushita. Research in the key areas of digital television systems and large-capacity data-storage devices is likely to see a significant boost this year.

The pharmaceutical industry is also considering more R&D spending. "Whatever the economic situation, we cannot afford to reduce our investment in R&D," says a spokesman for Eisai, one of Japan's largest pharmaceutical companies. **Asako Saegusa**

Mixed reaction to New Zealand reforms

[SYDNEY] A wide-ranging overhaul of university funding in New Zealand, announced last month in a long-delayed white paper on tertiary education, has pleased polytechnics and government-owned Crown Research Institutes, but attracted criticism from university leaders and researchers.

Wyatt Creech, the education minister and deputy prime minister in the National Party minority government, has watered down some controversial proposals from the 1997 green paper, such as turning universities into government-owned corporations (see *Nature* 392, 320; 1998).

Statutory independence is retained, although governing councils will be halved to a maximum of 12 members, with minimal academic representation and government "intervention" for "high-risk institutions".

Creech will retain overall financial control, but substantial funding will be shifted from universities to polytechnics and the growing number of Private Tertiary Establishments, including those from overseas.

Tony Steel, the National Party's chair of the parliamentary education and science committee, told last week's Association of University Staff conference that New Zealand has too many universities, and research funds "were going to some arts disciplines where it was questionable that the outcomes justified the cost".

There will be more research programmes for polytechnic staff to ensure greater equity across sectors. A quality assurance authority will be set up, with funding being contingent on quality in teaching and research.

The Association of Polytechnics described the white paper as "positive, constructive and sensible". Jim Doyle, the association's executive director, says a polytechnic student presently has access to assets of just NZ\$14,000 (US\$7,300), compared with NZ\$29,000 for a university student.



Gould: Among those critical of changes.

There are 57,237 equivalent full-time students in 25 polytechnics and 83,528 in seven universities, receiving subsidies of NZ\$402 million and NZ\$684 million, respectively. "Tuition subsidies will be provided on the same basis, no matter where they [the students] are studying", says Creech.

The 34,000 students in private tertiary establishments will become eligible for subsidies.

From an estimated NZ\$100 million for postgraduate training in universities, a "contestable pool" of NZ\$20 million will be established for "targeted programmes" (allowing the government to direct research). The remaining NZ\$80 million will go to institutions through top-ups of tuition subsidies to degrees and postgraduate programmes.

Creech promised that funding and legislative requirements for research will be reviewed in 2001 "with a view to transferring a further NZ\$60 million to the contestable pool".

Despite the watering down, the changes have still been criticized by senior university officials. Daryl Le Grew, vice-chancellor of the University of Canterbury, fears the effect on "a top-end research university". He predicts a legal challenge over "institutional ownership", and attacks the plan for "seeking the safety of mediocrity".

Alastair MacCormick, acting vice-chancellor of the University of Auckland, sees "threats to institutional autonomy and therefore academic freedom". This is echoed by Jane Kelsey, president-elect of the Association of University Staff, who is concerned that the changes "will take money away from the libraries, technology and staff".

Bryan Gould, vice-chancellor of Waikato University and chairman of the NZ Vice-Chancellors' Committee, describes the creation of the quality assurance authority as "unnecessarily heavy-handed and otiose". He says it is being set up "not because of any perceived deficiencies in the universities' own quality assurance procedures, but because of a [political] need to establish uniformity across the sector".

The committee says the transfer from universities will cut funding for their postgraduate research programmes by NZ\$7.5 million a year. There is also scepticism about extra long-term funding after Maurice Williamson, the science minister, berated scientists for drawing attention to his recent retreat from a 1996 commitment to boost gross expenditure on R&D from 0.5 per cent of gross domestic product to 0.8 per cent by 2010.

Universities claim that the 'portability' of subsidies will add uncertainty to course planning, staffing and allocation of infrastructure. The white paper says the changes will improve research quality, a claim disputed by Gould, who says "it is hard to imagine a more damaging recipe for destroying the New Zealand research effort".

He describes as a "hammer blow" to universities a "capital charge" of reducing tuition subsidies for students attending "asset-rich" universities, with the saving being passed to "poorer" institutions.

The Association of Crown Research Institutes has been campaigning against what it sees as an unfair allocation of funding to universities (see *Nature* 391, 834; 1998), as the institutes have had to cut staff this year. Ian Warrington, the association's president, welcomes the government's "forcing a reassessment of research qualities within existing tertiary institutions".

Warrington suggests that "a blurring of the use of funds within universities" will be mitigated by contestability. But Kelsey predicts that the increase in contestable research funding — and decrease in non-contestable funding — would "result in those who taught being alienated from doing research".

George Petersen, a biochemist at Otago University and president of the academy of the Royal Society of New Zealand, speaking in a personal capacity, believes the policies will result in "a large number of poorly funded institutions, a curious way of going about the improvement of tertiary educational quality". He attacks Creech for giving no details of how the research fund will be administered.

In contrast, John Campbell, a physicist at the University of Canterbury and prominent promoter of science, argues that universities "are getting what they deserve for past arrogance and aloofness".

Peter Pockley

UK minister says sorry for science past

[LONDON] The 'shadow' science minister in Britain's opposition Conservative party has expressed regret for cuts to science funding while the Conservatives were in power.

In a speech last week to the pressure group Save British Science, John Redwood, the opposition spokesman for trade and industry, said he welcomed the current government's increase in the science budget (see *Nature* 394, 209; 1998).

But Redwood conceded that the previous Conservative government had not left science in as good a state as it should have done. "We accept we have made mistakes in the past. We have changed," he said. New Conservative science policies would

emphasize factors such as the need to provide generous tax incentives for investment in research by private investors.

Redwood added that the Conservative party was attempting to learn from its mistakes following its disastrous failure in the last election, and was listening directly to public opinion about its policies.

He also criticized the government for "dressing up" the value of the increase in spending on science. Under the terms of the comprehensive spending review a further £700 million (US\$1,157 million) will go to science. But the 15 per cent real terms increase over 1998–99 baseline funding will be spread over three years.

Natasha Loder

Monsanto set to back down over 'terminator' gene?

[WASHINGTON] The US life sciences company Monsanto is considering delaying the introduction of its controversial germination control technology — tagged by critics as 'terminator' technology — in which crops are genetically modified to destroy their own seeds.

A decision is expected after a meeting today (10 December) between key Monsanto executives and Robert Shapiro, the chairman and chief executive.

The proposed technology has caused anger among farmers in developing countries, who would not be able to save new seeds for replanting if they chose to grow germination control crops. It was developed jointly by the US cotton seed producer Delta and Pine Land Co. — which was taken over recently by Monsanto — and the US Department of Agriculture.

One of the reasons for developing germination control was to protect the intellectual property rights of those developing the seed. A Monsanto working group has been considering how best to respond to the controversy, which has virtually blocked any short-term prospect of improving the company's public image — and that of genetically modified food in general — in developed and developing countries (see *Nature* 396, 397; 1998).

Some in the company argue that Monsanto

to should open a dialogue with its critics, and organize independent research into the implications of germination control technology for all the stakeholders — including the company. The company had previously been asked by signatories to the United Nations Biodiversity Convention to help with their own review of germination control. This exercise, however, has not yet begun.

But any moratorium on introducing germination control would cause internal problems for Monsanto, where such a move is the focus of a vigorous debate. Some staff believe that, far from improving the company's public image, a moratorium might have more damaging consequences in the long term.

They argue that it could be seen as a victory for Monsanto's critics against not just germination control, but the whole of agricultural biotechnology. Such critics include the Rural Advancement Foundation International, based in Canada, and the Research Foundation for Science, Technology and Natural Resource Policy in India.

These are campaigning for a worldwide ban. They achieved some success in October when germination control seeds were banned by the World Bank's agricultural research agency, the Consultative Group on International Agricultural Research. **Ehsan Masood**

Allow cloning in embryo research, says UK report

[LONDON] The use of cloning techniques should be permitted in human embryo research under restricted circumstances, recommends a report by the UK Human Genetics Advisory Commission (HGAC) and the Human Fertilization and Embryology Authority (HFEA). But the two groups strongly support the existing government ban on human cloning for reproductive purposes and suggest considering legislation to forbid this.

The report suggests cell nucleus replacement techniques, which "do not involve human reproductive cloning", hold promise for the treatment of serious illnesses, such as Parkinson's and Alzheimer's. The report follows a public consultation exercise launched in January. The HGAC was established in December 1996 as a non-statutory body providing advice to the health and industry secretaries on issues arising from developments in human genetics that have social, ethical and/or economic consequences.

The report, *Cloning Issues in Reproduction, Science and Medicine*, attempts to draw a fine line between two types of cloning in order to leave the door open for certain research to continue. The document distinguishes between "reproductive cloning or the production of an entire animal from a single cell by asexual reproduction — and therapeutic cloning, applications of nuclear replacement technology which do not involve the creation of genetically identical individuals".

Under existing rules, researchers would be unable to exploit avenues opened up by the recent breakthrough in the culture of human embryonic stem cells (see *Nature* 396, 104; 1998). The report would extend the current authorized use to include two further purposes for which the HFEA might issue licences for human embryo research — for the development of therapeutic treatments for diseased or damaged organs, and for the development of therapeutic treatments for mitochondrial disease. The report also suggests that, in the light of rapid developments in science and in public attitudes, the situation should be re-examined in five years.

The move follows the release last week of a report on public opinions on human cloning. This concluded that the public felt that humans should never be cloned for reproductive purposes and had reservations about the idea of using cloning technology to create stem cells to make tissues and organs for medical treatment. The Wellcome Trust report found that the public are able to understand the science and grasp the ethical issues and social implications. **Natasha Loder**

Brazil's scientists fear ministry merger

[SÃO PAULO] Brazil's Ministry of Science and Technology could be merged with the larger Ministry of Education when the second term of President Fernando Henrique Cardoso begins officially in the new year.

Brazilian scientists fear that the merger could result in reduced funding for research. The plan is one of three options in a report sent to Cardoso by his Minister of Education, Paulo Renato Souza, a former rector of the University of Campinas, who would probably head the resulting superministry.

The second option, which scientists are equally concerned about, is to give the science ministry responsibility for the 52 federal institutions of higher learning, including all federal universities and some independent centres. The third option would see the current arrangements unchanged.

The science ministry handles about half the total federal spending in research and development. It is currently responsible both for supporting research through grants agencies such as the National Council for Scientific and Technological Development,



Souza: tipped to head new superministry.

and for directly managing some research centres.

The difficult financial situation facing the federal universities is one reason for the possible changes. According to calculations by the National Association of Directors of Federal Institutions of Superior Learning, 94 per cent of

the 5.7 billion reais (US\$4.7 billion) that the Ministry of Education gave these institutions this year was spent on salaries.

Sérgio Henrique Ferreira, president of the Brazilian Society for the Development of Science, thinks that merging the ministries is "madness".

His concern is that it could mean less spending on research. He says both ministries have their own grants agencies, and the government may decide to limit financing of these, arguing that they are duplicating efforts. **Ricardo Bonalume Neto**

US scientists may boycott AIDS congress

[CAPE TOWN] Some researchers are considering boycotting the World AIDS Congress in Durban in 18 months' time, in protest at the decision by South Africa's health minister to abandon a planned pilot programme to administer the antiviral agent AZT to HIV-positive pregnant women.

Nkosazana Zuma's decision was announced in October after a meeting with the health ministers of the country's nine provinces, which are responsible for providing health services in South Africa's federal system. The reason given was that funds to support the programme were not available at provincial level.

The move sparked an outcry, and the volume of complaints increased last month when the government announced plans to spend R30 billion (US\$5 billion) upgrading the country's defence force. Now, several US AIDS researchers say they may not attend the congress unless the policy is changed.

"This is something some of us are actively discussing — we would prefer it didn't come to that, but it might be necessary to get a boycott organized," said one US-based researcher.

Clinical trials in Thailand earlier this year

indicated that administering AZT over the last four weeks of pregnancy and during labour (a 'short-course' regimen) more than halves the mother-to-child transmission of HIV.

Zuma's decision to axe the programme appears not to have been based on the costs of the pilot programme — the French-based International Solidarity Fund has since offered to pay for this — but on fears that it could lead all HIV-positive pregnant women to believe they have a right to the drug.

"There is not much point in running a pilot study unless you can implement its findings," says physician Ian Roberts, special adviser to Zuma. But Roberts indicates that the minister's decision is not necessarily irreversible. "Next year the provinces might have a budget for the programme, depending on the cost at which Glaxo is prepared to supply the drug," he adds.

In neighbouring Botswana, the government has already decided to make the drug available to all such women. The cost of implementing the scheme in South Africa, assuming the pilot programme is successful, is estimated at R80 million (US\$14 million). The country is believed to have an adult infection rate of 15 per cent, with the per-

centage of infected pregnant women varying from 16 to 27 between regions.

But Glenda Gray, co-director of the perinatal HIV research unit at the University of the Witwatersrand's Chris Hani Baragwanath Hospital, says that a study has shown that the short-course regimen is ultimately more cost-effective than treating children who could have escaped infection.

"The pilot study would have teased out difficulties concerned with implementation, such as securing drug supplies, ensuring safe distribution and dispensation, supplying milk in lieu of breast-feeding, and problems related to responses to intervention, teaching and counselling," says Gray. "Its implementation makes good sense — the only reason Zuma could have made this decision is that she was ill-informed."

Gutaaf Wolvaardt of the Medical Association of South Africa, who is chairman of the organizing committee for the World AIDS Congress in 2000, says he is unaware of any moves by researchers to respond to Zuma's decision by boycotting the congress. "I would question the logic of this, as the congress is not organized by the Minister of Health," he says.

Michael Cherry

Palaeontologists divided over 'stay at home' policy for fossils

[LONDON] A group of senior palaeontologists is proposing an international agreement that hominid fossils should not be moved from their country of origin without "compelling" scientific reasons.

In a motion passed earlier this year during its meeting in South Africa, the International Association for the Study of Human Palaeontology (IASHP) urged that requests to move such material must be backed by a demonstration that the investigation could not proceed "in the foreseeable future" in the country of origin.

Not all palaeontologists agree. Some feel that the words "foreseeable future" are too vague, and are concerned that increased restrictions on the movement of fossil hominids could hold back scientific investigation, as the country of origin may lack up-to-date research equipment.

But those keen to press the case for restricting access argue that this is needed to compel greater investment in museums in developing countries. "If you allow fossil hominids to be taken [abroad] every time you want to do a sophisticated analysis, countries are never going to develop," says Bernard Wood, of George Washington University in Washington DC.

The need for restrictions to be tightened was initially raised by Meave Leakey, of the National Museums of Kenya, following a



Leakey: seeks to protect 'priceless' fossils and says some sophisticated analyses may be unnecessary.

request that Turkana boy, a 1.6 million-year-old skeleton of *Homo erectus*, be lent to the University of Chicago for an exhibition.

Wood, who put forward the IASHP resolution, says: "We all know developing countries need money. But many are concerned this is not the way to raise it. If it is done this way, there is no pressure on governments to fund museums properly."

Wood adds that, if a government cannot find money for a scientific investigation,

"then it is up to the international community to solve [the problem], without the need to have fossils flying round the world in money-making activities".

But a senior British palaeontologist says that, although he welcomes the attention that the resolution has brought to the issue, he does not support the wording. "It is all very well for scientists to pass this, but it is still museums and governments who make decisions," he says.

"All of us are being continually asked to loan specimens," he adds. "Our policy 20 years ago was not to allow transportation. But in the last ten to fifteen years we have loosened our attitude."

He points out, for example, that the Natural History Museum has been unable to afford expensive computerized tomography and γ -ray dating facilities, and that some "outstanding science" has resulted from its loans to institutions abroad. "The science community is going to have to be realistic," he says.

Leakey describes hominid fossils as "priceless specimens". She says: "We are often asked very difficult questions — such as requests to apply destructive techniques. I ask, what are you going to gain? Technology progresses so fast, you have to be careful not to make a hasty decision."

Natasha Loder

Animal rights activists turn the screw

[MUNICH] Biomedical researchers in Germany and Britain are calling for more support from research organizations and politicians following a wave of violence and death threats by militant animal rights activists.

Mainstream animal rights organizations have condemned the activists' tactics. But in Germany they are pressing for a constitutional change to protect the rights of animals, a move that scientists fear could lead to lengthy delays in the approval of research protocols.

The British Animal Liberation Front (ALF), a militant activist group, threatened last week to kill ten British animal researchers and breeders if imprisoned ALF member Barry Horne, who has been on hunger strike for more than 50 days, should die.

Horne is serving an 18-year sentence for bomb attacks on the Isle of Wight and in Bristol. He is demanding a Royal Commission on animal experiments, which the Labour Party had promised before the 1997 general election, in addition to the existing committee on animal experiments.

Those receiving death threats include Colin Blakemore, a neurophysiologist at the University of Oxford and last year's president of the British Association for the Advancement of Science, who has already been attacked by animal rights groups. He studies the working of the cerebral cortex, but says that in recent years he has reduced the number of primates and cats used in his research by adopting tissue-culture techniques.

Several years ago, the Labour Party promised various changes to the British animal protection law. But since taking up government, it has proceeded more slowly than some critics would like. This, says Blakemore, is one reason for the recent activities by such groups.

In Frankfurt last week, angry protests accompanied the award of the Hessian Cultural Prize to Wolf Singer, a director of the Frankfurt-based Max Planck Institute for Brain Research and a former president of the European Neuroscience Association.

Regional animal rights groups wrote to the Hessian state government, referring to Singer as a "monster", a "non-person" and a "cultural disgrace". Singer, who has received several death threats, was given police protection during the award ceremony.

Andreas Kreiter, one of Singer's former students who studies the electrical properties of primate brains, has frequently been subjected to physical and verbal attacks since moving to Bremen nearly two years ago. He has been under permanent police protection since activists tried to force their way into his office last year, and now he works in an isolated laboratory under constant guard.

The hostilities against Singer and Kreiter



Pressure for change: supporters of UK hunger-striker Barry Horne want a royal commission.

were sparked by their experiments on the visual system of macaques. Singer's work on the functioning of the cerebral cortex is likely to contribute to future treatments of brain-related diseases, such as schizophrenia or Alzheimer's disease.

Even researchers who have not experienced hostile acts, such as Nikos Logothetis, a director at the Max Planck Institute for Biological Cybernetics in Tübingen, admit to feeling uneasy. Logothetis worked for 17 years as a neuroscientist in the United States before coming to Germany last year. "I hope this was not the wrong decision," he says.

Mainstream animal protectionists, such as the Deutscher Tierschutzbund, the German association for animal protection, and the British Union for the Abolition of Vivisection, say that animal research cannot be justified on scientific or medical grounds. Germany has the strictest animal protection law in the European Union (see *Nature* 391, 624; 1998), but animal rights groups are campaigning for a ban on animal research.

But they distance themselves from extremist and militant actions. Representatives of German and British animal protection groups describe violence as "embarrassing" and "counter-productive" to their goal of collaborating with researchers to develop alternative methods.

The Deutscher Tierschutzbund hopes the German government's plan to include animal rights in the constitution will help to reduce the number of animal experiments. A legislative initiative is likely to start soon, and the two-thirds majority required for a constitutional change is likely to be achieved, as the initiative has cross-party support.

This could have far-reaching consequences for animal researchers. The constitutionally guaranteed freedom to research would have to be weighed up against the constitutional rights of animals to be protected from avoidable pain. Thomas Schröder of the Deutscher Tierschutzbund says the association would immediately initiate court

cases and injunctions against researchers.

Martin Steins of the Max Planck Society (MPS) believes it would take at least ten years to clarify the legal position of researchers who use animals.

Jan-Erik Bohling, spokesman of the Society of Health and Research, Germany's main lobby group for academic and industrial research, says Germany would be isolated from research carried out elsewhere. "The approval procedure for animal experiments would become insanely drawn out," he says.

Many researchers feel abandoned by research organizations, such as the MPS and the Deutsche Forschungsgemeinschaft (DFG) and by their colleagues. Indeed, some faculty members at the University of Bremen have issued a memorandum distancing themselves from Kreiter.

Blakemore argues that "medical and scientific institutions have not been public enough in their support of animal research". He has frequently appeared on television and in other high-profile media to explain to the public the value of his work, exposing himself to risk of attack in so doing.

Last month, Singer detailed the scientific, medical and ethical justifications of animal research in an article published in the *Frankfurter Allgemeine Zeitung*.

Hubert Markl, president of the MPS, last week strongly defended Singer against the "malicious campaign" of his extremist opponents. But both the MPS and the DFG, aware of public sensibilities, are reluctant to fight openly to defend animal researchers.

"A constitutional change would not be necessary if researchers limited their animal experiments to those that are really important," says Reinhard Grunwald, secretary general of the DFG. "After all, the DFG is not supposed to act as a lobby group for the use of animals in research."

But most of the animal researchers affected stress that they are not going to admit defeat. "I am not going to give way to blackmail," says Blakemore.

Quirin Schiermeier

Lawyers review US ban on stem cell research funds

[WASHINGTON] Federal funding for research into human embryonic stem cells will remain prohibited in the United States pending a review by lawyers from the National Institutes of Health (NIH) on the correct interpretation of the current law banning such funding, Harold Varmus, the NIH director, told a Senate committee hearing last week.

The Senate appropriations committee for labour, health and education, chaired by Senator Arlen Specter (Pennsylvania, Republican), appeared inclined to revise the current ban. This followed evidence from Varmus and others about the potential of the stem-cell research to revolutionize the treatment of disease by enabling the growth of the desired type of body tissue.

But experts testifying at the hearing on 2 December pointed out that the research that isolated the stem cells, which was privately funded by Geron, a Californian biotechnology corporation, may have been permissible under the current law anyway, as they were isolated from fetal tissue, rather than from embryos.

Varmus says the NIH will review its

interpretation of the law before supporting any further research on the stem cells. An NIH spokesman said it was not known when this review would be complete.

German science minister seeks extra billions

[MUNICH] Germany's new Social Democrat minister for education and research, Edelgard Bulmahn, is to press the government for an increase of DM1 billion (US\$600 million) in her budget each year until 2003. The extra money would be spent on increasing student grants, promoting the careers of young scientists, and key research areas.

Bulmahn defines the latter as including environmental protection, health, genetics and biotechnology, transport and communication technology. But she has toned down her party's pre-election promises to double research spending in the next five years by saying that the promise referred only to money for research programmes, not to basic institutional funding.

Rising tide of threats to marine ecosystems

[WASHINGTON] Marine ecosystems on the eastern coast of the Americas are facing an upsurge of environmental problems,

including harmful algal blooms, diseases to marine organisms, and human diseases borne by marine bacteria. So says a report published this week by the Center of Health and the Global Environment at Harvard Medical School.

The study, led by Paul Epstein and funded by the National Oceanic and Atmospheric Administration, finds that these phenomena have increased dramatically in marine ecosystems along the eastern coast from Canada to Venezuela over the past 25 years. Epstein blames human activity, including pollution, loss of wetlands and overfishing, for the trend.

Company shut down over nuclear fuel scandal

[TOKYO] The Japanese nuclear-power engineering company that was discovered earlier this year to have falsified data on 40 of 44 containers used to transport spent nuclear fuel, announced last week that it will voluntarily dissolve next summer to take responsibility for the action.

The decision was announced during a news conference held jointly with Japan Atomic Power, a government-run organization with a 100 per cent stake in the company, Genden Engineering Services and Construction Co. In addition to the

resignation of the company's executives, those from four major electric-power firms and officials from related ministries — including the director of the Science and Technology Agency's nuclear-safety bureau — have received official warnings and will have their salaries cut by up to 50 per cent.

British biotechnology to escape more red tape

[LONDON] The UK government is resisting pressure to impose greater regulation on the biotechnology industry in the wake of accusations that companies are exaggerating the potential of their products to investors.

Responding to a recent inquiry into the company British Biotech by the House of Commons Select Committee on Science and Technology, the government said it "does not see a need to seek further changes to the medicines licensing system". But a cabinet committee on biotechnology will review a range of issues on regulation.

Plan to relocate Paris universities stalled

[PARIS] The future location of some 10,000 Parisian researchers is hanging in the balance, after an interministerial meeting last week postponed a decision on the

possible relocation of the universities and institutes on the sprawling Jussieu campus in the Latin quarter. The campus is scheduled to undergo a FF4 billion (US\$ 0.7 billion) programme to remove asbestos (see *Nature* 382, 291; 1996).

Claude Allègre, the science minister, recently lent his support to the proposed transfer of one of these institutions — the University of Paris VII — to a development zone beside the new Library of France in south-east Paris. But the finance ministry seems in no hurry to agree to a final decision. It is faced with the costs of the asbestos removal, and the desire of several other universities to consolidate their fragmented activities at the same site.

Brussels funds natural history museum grants

[LONDON] Natural history museums in Paris and Amsterdam have been granted Large Scale Facility status by the European Union. This means that researchers from European and associated states can apply for funding from the European Commission for visits to the museums of up to three months.

The move follows a decision to offer this status to the Natural History Museum in London earlier this year — said to be the first time a museum has been given this status. All

three museums are now jointly seeking applications from European researchers interested in using their facilities.

World science congress to talk equity and ethics

[SYDNEY] Equity in access to science for women, how to manage the marine environment, assuring supplies of water and food, and unambiguous standards for the ethical behaviour of scientists should be high among the priorities for next year's World Conference of Science in Budapest, according to delegates from 27 Asia-Pacific countries who attended a meeting in Sydney last week.

NASA mission targets Mars climate studies

[WASHINGTON] The US space agency NASA's next mission to Mars, the Mars Climate Orbiter, is scheduled for launch today (10 December) from Cape Canaveral, Florida, on a Delta rocket.

Carrying an infrared radiometer and a colour camera, the spacecraft will arrive at Mars next September to study the planet's atmosphere. A separate Mars Polar Lander will be launched on 3 January, and will land near the Martian south pole the following December.

No room for complacency over climate

Sir— In their Commentary on the prospects of future climate change, Martin Parry *et al.* state that the Kyoto Protocol “is an agreement to a 5.2 per cent reduction in greenhouse-gas emissions by about 2010 (relative to 1990), and constant emissions thereafter” (*Nature* 395, 741; 1998). Any reader of the actual text will see that it says no such thing (<http://www.unfccc.de>).

The protocol — like its predecessor, the United Nations Framework Convention on Climate Change — sensibly takes a step-by-step approach, setting targets only for the budget period 2008–2012. What countries choose to set as future commitments, and who will be involved in those commitments, is a matter for future negotiation. Periodic review of commitments is required under the convention and the protocol. The protocol commits parties to agreeing commitments for future budget periods by 2005 at the latest.

Regrettably, therefore, some of the subsequent argument in the Commentary breaks down. It is, however, a reminder of what would happen if countries were complacent in those subsequent review periods, and a warning that even the most ambitious steps will leave some parts of

the world with serious adaptation issues.

David Fisk
(Chief scientist)

Department of the Environment, Transport and the Regions, 123 Victoria St, London SW1E 6DE, UK

Parry *et al.* reply— David Fisk is correct in drawing attention to our interpretation of the Kyoto Protocol. Because the protocol sets targets only through to 2008–2012, we assumed that emissions would be constant thereafter, but only as a starting point for our analysis. We recognize, and indeed hope, that further reductions in emissions beyond 2012 will be agreed under the protocol, and we gave estimates of impacts

Table 1 Estimates of global warming for the year 2050

Emissions scenario	Global warming (°C) with respect to 1961–90
Unmitigated	1.39
Kyoto	1.33
Kyoto+	1.29
Kyoto++	1.24

Kyoto+ assumes annex I nations continue to reduce emissions over the period 2012–2050 at 1 per cent per year (similar to their reduction over the period 1998–2012).

Kyoto++ assumes the above, plus non-annex I nations reducing their emissions over the period 2020–2050 by 1 per cent per year.

following substantially greater reductions.

Contrary to what Fisk states, our argument about the necessity for adaptation remains intact. Indeed, under two more progressive assumptions about post-2012 emissions targets, global warming by 2050 — and the associated impacts — will still be substantial (Table 1). The assumption of continued reductions in 38 industrialized ‘annex I’ countries, and the involvement of non-annex I nations by 2020, reduces warming by the year 2050 by only 0.15 °C. Such post-2012 commitments will of course yield larger benefits in the longer term, but the inertia in the carbon cycle and climate system means that we will need to ‘adapt to the inevitable’ in the medium term.

Martin Parry, Matt Livermore

Jackson Environment Institute, University College London, 5 Gower Street, London WC1E 6HA, UK

Mike Hulme

Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK

Nigel Arnell

Department of Geography, University of Southampton, Southampton SO17 1BJ, UK

Robert Nicholls

Flood Hazard Research Centre, Middlesex University, Enfield EN3 4SF, UK

Open up review system

Sir— The issues of peer review and authorship are of central importance to scientists, and are frequently debated in journals. But many such discussions fail to refer to the increasing number of critical investigations into the workings of the publishing process.

The 15 July issue of the *Journal of The American Medical Association*, for example, contains 33 articles based on presentations at the 3rd Congress of Biomedical Peer Review in Prague in 1997, and constitutes a rich source of analyses of current and proposed publication procedures in biomedical research. Such studies suggest that two easily implemented systems could improve current practice.

Listing authors’ contributions, and listing those who could guarantee the integrity of an article, offer a big improvement over the current ways of dealing with authorship. Similarly, the predominant system of editorial review, where the reviewers are unknown to the authors, is considered unfair. The two justifiable systems are a fully closed one (where the reviewers, authors and ideally the editors are unaware of each others’ identities) and a fully open one (where all

the parties know each others’ identities). Research shows that the fully closed system suffers from poor success in masking identities, suggesting that the fully open system is the more favourable alternative.

It is encouraging that both explicit listing of contributions and open peer review are starting to be adopted (for example in *The Lancet* and *Cell Calcium*, respectively). But faster and more general implementation of these systems would be facilitated by a leading interdisciplinary journal such as *Nature* showing the way.

Bostjan Kobe

St Vincent’s Institute of Medical Research, 41 Victoria Parade, Fitzroy, Victoria 3065, Australia

Whose law for sharing research tools?

Sir— Your editorial on the sharing of research tools called for a uniform materials transfer agreement (MTA) (*Nature* 396, 97; 1998). As a research scientist, I endorse this proposal, but there is a problem that requires urgent attention.

The legal security that MTAs offer to those who sign them requires a legal basis — but which law should apply? MTAs

issued by the US National Institutes of Health (NIH) require that US law applies. The NIH also insists that anybody providing it with material signs an indemnity clause in its favour. This means that my institution would be liable for damages resulting from the use of any research material that I sent to the NIH.

Anyone who provides research tools should take responsibility for their safety. However, they must be allowed to express the reservation that the material might be partially unknown and potentially hazardous, so removing liability for incidents that might result from its use in experiments. It also worries my research institution, which cannot accept the NIH’s MTAs, that US law should apply to material we send to them. We insist that German law applies, because donors must be in a position to assess the legal consequences if they give sensitive material to others. This is particularly important given the US enthusiasm for litigation.

Solving this problem should be a challenge for the lawyers, the enjoyment of which might even compensate for their loss of business if MTAs are made simpler.

Rudi Balling

Institute of Mammalian Genetics, GSF Research Centre of Environment and Health, Ingolstädter Landstr. 1, 85758 Neuherberg, Germany

Relativity and relativism: who's confused?

Sir — In his review of the book *Intellectual Impostures*, Richard Dawkins makes the astonishing claim that the authors, Alan Sokal and Jean Bricmont, “expose Bruno Latour’s confusion of relativity with relativism” (*Nature* 394, 141–143; 1998).

Here is what Sokal and Bricmont say about this matter in their book: “Finally, Latour draws an eminently sensible distinction between ‘relativism’ and ‘relativity’: in the former, points of view are subjective and irreconcilable; in the latter, space-time coordinates can be transformed unambiguously between reference frames.”

Perhaps, after Dawkins explains how “Latour draws an eminently sensible distinction” became “Latour’s confusion”, he will tell us why he feels entitled to claim that there is a confusion in an essay of Latour he shows no sign of having read. He could hardly have done so without noticing that Latour, in fact, belabours the distinction between relativity = objectivity (good) and relativism = no objectivity (bad). Is Dawkins really talking about Bruno Latour?

The review is shot through with misinformation. I believe Dawkins is sincere. But he has failed badly to take the measure of the texts and people he talks about, both those he favours and those he does not. For example, an unhurried reading of page 8 of *Strange Weather* shows that he badly misunderstands Andrew Ross’s ironic reference to how his lack of scientific training (“the science teachers I never had”) determined what book he wrote (“it could only have been written without them”). Referring to arguments about the cultural authority of intellectuals that he developed in an earlier book, Ross explains: “My plan was to explore the reach of these arguments into the world of science and technology. As I lacked the training of a scientific intellectual and the accompanying faith, however vestigial or self-critical, in the certainties of the scientific method, it became clear that my point of identification could not be with ‘high’ scientific culture. My position, then, became that of a cultural critic examining the power and authority of the claims made for science and technology [and] the responses to these claims in the popular culture.”

An equally attentive reading of page 113 of Sandra Harding’s *The Science Question in Feminism* refutes the canard repeated by Dawkins about “a notorious feminist description of Newton’s *Principia*” as a rape manual. The sexual metaphor in question — forced penetration of a female entity — is not twentieth-century feminist but, on

Harding’s reading of the literature, early modern scientific. She believes it probably was fruitful for early modern scientific enquiry and conceptualization, hence her ironic observation that “rape manual” seems as apt a metaphor for the early modern conception of Newton’s laws as is “mechanics”.

Dawkins would like works such as *Intellectual Impostures* to reach a wider audience. So would I. But, unlike Dawkins, I want them to be read sceptically and held to the strict standards of evidence and logic that he and the authors of these works invoke even as they abuse them. By these standards, “it makes me laugh” and “I don’t get it” are not arguments. *Nature* should start enforcing these standards too.

“Science as a cultural construct”, and the letters it generated, came close to meeting these standards (*Nature* 386, 545–547; 1997). This made it possible for important issues to be discussed in a thoughtful way. The same cannot be said of Dawkins’ review or the works that he recommends in it.

Gabriel Stolzenberg

Northeastern University, Boston,
Massachusetts 02115, USA

Dawkins replies — It is surely only in a postmodern climate that anyone could seriously feel the need to distinguish relativity from relativism. But, yes, Latour does indeed state the obvious explicitly. The trouble is that the good work is then all undone by his stated intention to show that “relativity itself could be said to be social”, and his reading of Einstein’s text as “a contribution to the sociology of delegation” (Sokal and Bricmont, page 115).

I’m sorry if I misunderstood Latour, Ross and Harding. But, for heaven’s sake, isn’t being misunderstood precisely what these people do for a living?

Richard Dawkins

Oxford University Museum of Natural History,
Parks Road, Oxford OX1 3PW, UK

Put a match to pyre review

Sir — Once again, you have not hesitated to expose my ‘errors’ to your readers, devoting significant space to beating the drum about my second Ig Nobel prize, awarded by ignorant self-appointed guardians of the purity of science (*Nature* 395, 535; 1998). But your readers deserve an opportunity to judge for themselves our results and their theoretical consequences.

I refer them to our website, which contains a simple protocol for those who wish to duplicate our experiments (<http://www.digibio.com>). Put briefly, the

current short-range electrostatic theory of molecule interaction–recognition via random collision cannot help us understand how biological reactions really work. By contrast, the presence in water of long-range electromagnetic fields, as proposed by quantum electrodynamics, sheds light on key features of biological systems. We can now record specific electromagnetic signals.

Isn’t it time to open the door to genuine scientific debate? Based on my own painful ten-year experience, we may as well start by throwing out the peer- (or is it pyre?) review system, which has become, behind a facade of excellence, the main antibody blocking the nearly deceased corpus called scientific free exchange, which once was the cornerstone of progress. It condemns any advance that is “hard to reconcile with what we know...” (Dudley Herschbach), a position which is the negation of scientific research. Surely this attitude, reflected in the scientific press, is the clearest indicator of the urgent need for uncensored media.

Jacques Benveniste

Digital Biology Laboratory,
32 rue des Carnets, F-92140 Clamart, France

Pull off those price tags

Sir — As they try to dispel the “myth of market prices”, Don Fullerton and Robert Stavins argue that economists are not exclusively concerned with the financial value of things (“How economists see the environment”, *Nature* 395, 433–434; 1998).

They say, for example, that the value of damages to health from environmental pollution not only includes healthcare costs and wage losses, but also “pain and suffering”. The goal of economic analysis being to measure the total value of the loss that individuals incur, “economists insist on trying to convert all these disparate values into monetary terms because a common unit of measure is needed to be able to add them up”.

But, unfortunately, neither a disappeared species nor a deceased beloved one can be bought back to life. What is the point of pricing things that no one can buy or sell? Doing so is erroneous and dangerous. It is easily understandable how the need for rational policies leads to the kind of analysis described by Fullerton and Stavins. However, the world does not fit entirely in the concepts of economic science, and sometimes rational decisions are simply not possible. In those cases, clearly stated moral principles, not shaky science, should guide decision-makers.

Vincent Detours

988 Acequia Madre, Santa Fe,
New Mexico 87501, USA

The world needs more than protests

The green movement has an impressive history of drawing attention to environmental problems. But if these are to be solved it must be more closely engaged in forging partnerships with business and government.

Pete Wilkinson

Over the last three decades, Green protest groups and organizations have legitimized the concept of environmentalism. Initially ridiculed and dismissed, they have stamped their mark on the international scene, bringing environmentalism into the common lexicon — and even making it fashionable.

At the same time, however, the problems highlighted by these groups have been getting worse. In October, for example, the *Living Planet* report, produced by the World Wide Fund for Nature International, the New Economics Foundation and the World Conservation Monitoring Centre, claimed that in the past 30 years we have destroyed 30 per cent of the natural world, with crippling impacts on forests, freshwater and marine ecosystems.

Having alerted the world to its plight, therefore, it is important that such groups now capitalize on their achievements by demonstrating that they can say 'yes' with as much conviction and determination as they have historically said 'no'. The whistle-blowers of recent years must learn to be arbitrators, conciliators and negotiators, catalyzing action from governments, industry, scientists and communities.

If environmental legislation is the stick for such action, then employment, business opportunities, greater equality and a more sustainable environment is the carrot. The same single-mindedness that environmentalist groups have shown in the past must now be used to turn the politics of protest into programmes of innovative cooperation.

Green groups must do more than campaign if they are to trigger solutions to the problems they draw attention to. Whistle-blowers will always be needed to expose recalcitrant, uncaring, or irresponsible companies and governments. But the need now is for projects through which we can demonstrate our global commitment to sustainable development.

This requires a new definition of environment. Whatever the limitations of the 1992 Earth Summit in Rio de Janeiro, one of its achievements was to give the concept of 'environment' a broader and more inclusive definition. What is the point of saving the sea turtle, for example, if the social, economic and political circumstances which have driven the species to the brink of extinction are left unaddressed?

This wider — and long overdue — interpretation encourages cooperation between stakeholders, and both broadens and



Wrong target? Brent Spar protests may have diverted attention from more pressing issues.

strengthens the green platform. It also recognizes that environmental damage arises from a complicated matrix of conditions and determinants; it does not start, or even finish, with the outlet pipe or the harpoon.

Patterns of consumption

We must, for example, relate environmental problems to the continually widening gap between rich and poor. This gap is itself partly the result of patterns of consumption which have fuelled resource depletion, habitat destruction and economic imbalance. The global movement of capital creates unemployment and impoverishes workforces as it shifts industrial production to labour-rich countries that often have a more cavalier attitude towards protecting the environment.

In this process ecosystems are destroyed, children die from environmentally-linked causes, water supplies become unusable. At the same time, consumers in rich countries eat smoked salmon from migratory fish raised in pens and doused in parasite-repelling chemicals, while poor artisan fishermen are forced to use dynamite and cyanide to kill ever-smaller fish to feed their families.

This gloomy picture can no longer be dismissed as the doom-laden rhetoric of a green lobby eager for attention. Furthermore, it

also contains a threat to social stability. An editorial in the *International Herald Tribune* in 1996 warned that "unless serious corrective action is taken soon, the backlash could turn into open political revolt and destabilize the western societies". If such words had come from green campaigners, they would have raised few eyebrows. But they came from the pens of two senior representatives of the World Economic Forum, whose members include the world's 1,000 largest global corporations.

Time, in other words, has run out on the environmental movement's efforts to realize the utopian dreams of the 1960s. Today we should not be continuing merely to point accusing fingers at wrong-doers, but expediting solutions collectively and cooperatively. Responsibility for achieving this does not fall to the greens alone. But their experience and public support has left them better equipped than many other groups to perform a pivotal role at the heart of societal reconstruction in the new millennium. For whatever the diversionary impact of their recent tactics, their effectiveness remains awesome.

In the Brent Spar incident of 1995, green pressure led the Shell oil company to abandon plans to sink an oil storage platform in the North Atlantic. Since then, European green protest groups have invested much

time and money ensuring that redundant oil and gas platforms are brought ashore for recycling rather than dumped at sea. Yet the wisdom of that decision — carried overwhelmingly by the Oslo and Paris Commission last summer — has not been questioned publicly. Issues related to the disposal of platforms, such as energy balance, landfill impacts, employment, carbon dioxide generation, artificial reef creation, establishing 'no fish' zones, or using rigs to delineate no-go areas, have all been subsumed under a roar of misplaced triumphalism.

In the United States, the principle that oil companies should fund environmental initiatives with the money saved by toppling *in situ* rather than recycling redundant rigs is becoming increasingly accepted. Yet the most powerful country in the world equates a reduction in carbon emissions only with lower energy demand, and thus with all the attendant horrors this purports to bring, such as market intervention, the curbing of incentive and growth, and the lessening of 'opportunity'.

It is now widely accepted that only a uniform reduction of about 30 per cent in global carbon dioxide emissions will have any significant impact on mitigating the effects of man-made climate change. Yet in practice, even the commitment at the Kyoto and Buenos Aires climate meetings to a cut-back of only 6 per cent required monumental bouts of international negotiating over a matter which affects all our futures. If only the green zealotry which has forced the oil industry to spend billions on one side of the Atlantic on a programme which may have only marginal environmental benefits could be unleashed on the other side on those who appear not to have grasped the full meaning of the precautionary principle.

Power of positive thinking

It isn't all gloom. Some members of the green movement are adapting in significant ways, forging links between business, politics and environmentalism. The Greenpeace CFC-free refrigerator is a good example. Other welcome examples of innovative breakthroughs are agreements between the Environmental Defense Fund and McDonald's, partnerships to secure carbon fixing arrangements, the Body Shop and its diversification into wind-generator technology, the Forest Stewardship Council and its sister initiative the Marine Stewardship Council, round-tabling and conflict resolution. And of course, there are abundant examples of successful projects throughout the world.

Lasting and fundamental success, however, generally comes in small parcels. More often than not this means smaller, largely unsung organizations working with local communities, traders and governments to deliver a sustainable and healthier lifestyle. Small scale, low- and appropriate-techno-

We must relate environmental problems to the continually widening gap between rich and poor

logical improvements to peoples' lives all over the world — the vital work of the aid and humanitarian agencies — are saving or enhancing the lives of tens of thousands.

On the broader screen the picture is bleaker. The environmental crisis that we discuss in relative comfort in the industrialized world is a living reality for millions. Never before has incisive activity been so urgently needed to cut a swathe through a decade of hype, half-truths and posturing. We must demonstrate that it is possible, desirable, and profitable to beat the sword of traditional green protest into the ploughshare of sustainable lifestyles.

The green lobby has already seen a dramatic widening of its ranks. In Germany, they have translated protest into a political movement which has seen them become the third most popular political party. They now hold the balance of power, placing them in a position most green protest groups would die for: they can effectively end the nuclear industry in Germany, and will almost certainly call a halt to the sending of spent nuclear fuel to the UK and France for reprocessing, something that 30 years of protest has failed to achieve.

But there has also been increasing polarization. Supporters of Earth First and animal 'liberation' have become more and more uncompromising in the face of the perceived acquiescence of green protest groups. Indeed, unless we aim to instigate a paradigm shift, moving public and political action swiftly in the direction of implementing the solutions to the problems we face, the ground so painstakingly won is in danger of being eroded completely by the powerful forces of environmental conservatism.

What we now need is a visionary and public alliance: non-governmental agencies



The way forward? The World Land Trust's La Milpa Field Station in Belize.

from all disciplines should pool a proportion of their income to create a joint fund committed to forging links with the major players in the corporate sector. One body which comes close to matching this template is the British-based charity World Land Trust. In the Philippines, it established a project on a biologically important island, bought ahead of developers who wanted to turn it into a luxury resort. The trust funded its purchase with loans generated by a business plan based on government-recognized protected fishing zones, a sustainable development programme, participation from the local community, saving the island's biodiversity and cooperation from the regional and state political apparatus.

Redefining environmentalism

Even more impressive is the trust's work in Belize, where organizations it helped to establish own and sustainably manage 230,000 acres of land. In a text-book project, the Programme for Belize works with the European Union, the Belizean government, the Massachusetts Audubon Society, the Nature Conservancy, local communities, university departments, industry and local non-governmental organizations to protect and sustainably manage a large tract of land, packed with wildlife, which would otherwise have been sold off to forest-raiding Menonite farmers.

Over four years, the trust has developed a low-impact resource identification/land management procedure jointly with the United States. In only its first full commercially operational year, the Belize project got nearly double the normal price for mahogany products certified as sustainably produced by the Forestry Stewardship Council. The greatest criticism of the trust's activities is its near-invisibility. But as John Burton, its Executive Director, disarmingly comments: "We are too busy saving land to spend time on publicity."

Armed with this new definition of environmentalism, we must now develop projects which put it into practice. It is imperative to interpret and communicate a concept of global citizenship in novel, visionary projects which fire the imagination. Commercialism and environmentalism can and must co-exist, and social entrepreneurialism is both acceptable and necessary.

In short, the way forward is through a union of all the major stakeholders, using governmental partners to ensure legislative backing for their actions. Where that is already happening, it reinforces a belief that a reinvigorated and refocused green movement can indeed change the world.

Pete Wilkinson, former director of Greenpeace International, is now an environmental consultant. He can be contacted at Blyth House, Bridge Street, Halesworth, Suffolk IP19 2AB, UK. e-mail: wilks@btinternet.com

Twenty-eight ways to build a solar system

Brett Gladman

Ever since their discovery, extrasolar planets have challenged standard ideas about planet formation. Faced with increasing numbers of 'unusual' planetary systems, astronomers are now developing models that attempt to explain these giant planets around Sun-like stars.

A bewildering variety of extrasolar planetary systems are now known. Although the systems that have been discovered are heavily biased (by the methods used to detect them) towards planets at least as large as Jupiter very near their stars, we are already glimpsing the diversity of solar systems¹. Some have Jupiter-like planets in circular orbits both near (closer than Mercury) and far (roughly the same as Jupiter) from their stars. Some planets have highly elliptical orbits, perhaps suggesting gravitational interactions with other large planets, although there are alternative explanations. It is believed that these newly discovered planets are 'gas giants', meaning that although they may have a solid core, most of the mass consists of hydrogen and helium. Writing in the *Astronomical Journal*, Levison and colleagues² have modelled a variety of possible 'outer solar systems'.

One general scheme for giant-planet formation consists of forming a set of 'embryos' (Mars- to Earth-size) by the aggregation of kilometre-sized 'planetesimals' from the disk-like nebula that swirls around a young star. The planetesimals form out of whatever solid material can locally condense from the nebula, depending largely on the local temperature. The embryos then come together to build giant-planet 'cores' with masses of 5–20 Earth masses; only then can cores suck up vast quantities of nebular gas and grow into gas giant planets³ before the remaining nebular gas is dispersed when the central star lights up.

If this description seems vague, then the reader has caught on. The physics of these processes is very poorly known. It is not really understood how planetesimals form or what the physical conditions (temperature, mass and density as a function of solar distance) of the nebulae are. Other mysteries include how long gas disks remain around forming stars before being blown away, how planets accrete gas, or how the presence of all that gas in the disks affects the orbits of the embryos. There are, of course, reasonable guesses.

Levison *et al.*² have made a brave attempt to study the diversity of possible planetary

systems by performing 28 numerical simulations of the aggregation of planetary embryos. Because the important physical parameters are unknown, they chose ranges that probably encompass the 'real' values. After constructing plausible initial distributions of embryos surrounding the star, they compute the orbital evolution of these protoplanets as they interact with each other, sometimes coalescing and sometimes throwing each other out of the system. Of particular importance (or perhaps concern) is the modelling of huge gas cocoons around the embryos to make them coalesce more

easily as they pass each other. This was done to prevent the known problem⁴ of the strong gravity of the protoplanets sling-shooting each other to high-velocity orbits and then bashing each other to bits before they get large enough to accrete gas.

Simulations continue until the orbits stabilize, and produce a wide variety of final planetary systems (Fig. 1). Some have a few widely spaced planets, others have many Uranus-sized planets evenly spaced, while yet others have big planets closely packed together. Some even look like our Solar System. Interestingly, the simulations do not produce the most prevalent class of known extrasolar planets — Jupiter-sized objects very close to their stars — indicating that some crucial piece of physics is missing. The authors observe that even the higher growth rates they use (produced by the gas cocoons) do not sufficiently damp the planetary velocities to produce solar systems like our own with nearly circular planetary orbits. This problem is also observed in simulations of inner-planet formation, as discussed in News and Views last week⁵, but because most of the accumulation is presumed to occur after the gas has vanished, it is not clear that gas is the culprit in removing the dynamical excitation. Apparently, some other damping physics must be operating, which many (including the authors) speculate comes from interactions with the hordes of smaller planetesimals not included in the simulations.

The variety found in these simulations raises two issues. One is that the range of physical parameters used (due largely to our ignorance) is far wider than the Universe actually produces in protoplanetary nebulae, implying that the diversity found is larger than would be observed in reality. For example, the authors double the gas mass of large cores on a timescale that varies by four orders of magnitude; perhaps real disks have much less variation, producing a smaller range of final planetary masses than emerge from these models.

The second issue is that of predictability for theories of solar system formation. The dynamical evolution of a small number of embryos is highly chaotic, giving variable locations and numbers for the resulting planets. Levison *et al.* find stable systems ranging from one to seven planets. So it is unreasonable to demand that planet-formation theories produce our Solar System in detail; rather, some of the time at least, they must provide systems similar to ours from what are thought to be plausible starting conditions. This is in contrast to more 'regularist' thinking of 10–20 years ago, when it was popular to postulate that most solar systems would have a Jupiter analogue just outside the 'snow line'. This is the distance beyond which the temperature has dropped enough that icy solids condense, and greatly increase the surface mass-density of solids and thus local accretion rates.

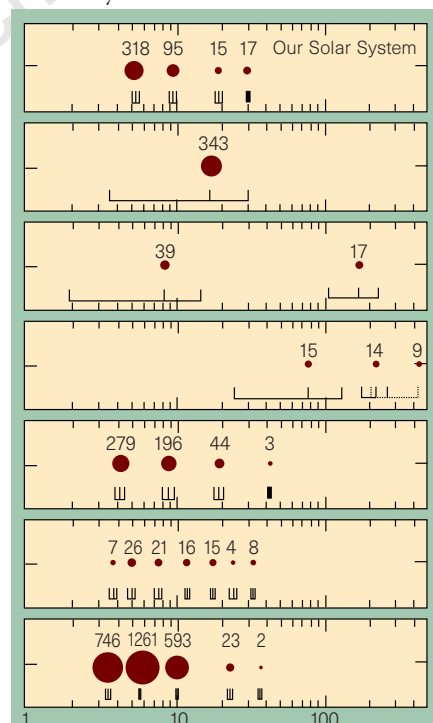


Figure 1 Some of the outer solar systems produced by the models of Levison *et al.*². The horizontal axis shows the distance of planets from the star, in astronomical units (the Earth–Sun distance), and the markings below the planets show the range of distances covered by each planet during its eccentric orbit. Numbers above the planets indicate their masses (in Earth masses). (Modified from ref. 2.)

The chaotic and violent character of this process, with planets throwing each other around and ejecting some members, is not good news for proponents of 'empirical distance relations', such as the Titius–Bode 'law' of planetary distances of our Solar System. If the search for stability is anything like that described in these models, then the arrangement of final solar systems is determined by a basically random process, subject to the final planets being "not too close"⁶, and meaning that a standard empirical law remains the dream of numerologists.

The near future holds the promise of a doubling or tripling of the number of extra-solar planetary systems known, with reasonable chances of detecting systems with planets both smaller and further from their stars. We may find that systems like ours are rare. However, progress in understanding the

physics of formation is most likely to come from the eventual study of protoplanetary disks around other stars on length scales of 1 astronomical unit — the distance from the Earth to the Sun. We may need to wait 5 to 15 years for optical interferometry and powerful coronagraphs, but if we want this information, it will come. □

Brett Gladman is at the Observatoire de la Côte d'Azur, Département Cassini, BP 4229, 06304 Nice Cedex 04, France.

e-mail: gladman@obs-nice.fr

1. Marcy, G. W. & Butler, R. P. *Annu. Rev. Astron. Astrophys.* **36**, 57–97 (1998).
2. Levison, H., Lissauer, J. & Duncan, M. *Astron. J.* **116**, 1998–2014 (1998).
3. Pollack, J. *et al.* *Icarus* **124**, 62–85 (1996).
4. Wetherill, G. W. in *The Formation and Evolution of Planetary Systems* (eds Weaver, H. & Danly, L.) 1–30 (Cambridge Univ. Press, 1989).
5. Hamilton, D. *Nature* **396**, 413 (1998).
6. Hayes, W. & Tremaine, S. *Icarus* **136**, 549–557 (1998).

Membrane fusion

Ready ... aim ... fire!

Randy Schekman

All cells use a programme of membrane fusion and fission to assemble membranes, both internally and on their surface. Given that biological membranes are essentially two-dimensional fluids, fusion must obey certain restrictions that prevent incompatible membranes from intermixing. Without such selective contact, membrane fusion would quickly lead to the loss of compartmental boundaries.

A powerful and unifying concept, the SNARE hypothesis suggests that membrane selectivity is achieved by the interaction of unique membrane receptors. The so-called v- (for donor vesicle) and t- (for target membrane) SNAREs join, and may even catalyse the fusion of, membranes that are designed to mix^{1,2}. But on pages 543 and 575 of this issue, Ungermann *et al.*³ and Peters and Mayer⁴ define three stages in the pathway — tethering (ready...), docking (...aim...) and fusion (...fire!) — that may challenge the standard model.

Many different v- and t-SNAREs have been described and, for the most part, each governs a unique step in membrane traffic. Genetic, pharmacological and biochemical data overwhelmingly support the idea that SNAREs are essential to fusion. Indeed, excellent progress has been made in solving the structure of SNARE proteins. Yet their exact function remains uncertain. For example, SNAREs implicated in diverse fusion reactions can form v–t complexes in solution that probably never exist in the cell. Moreover, the fusion reaction catalysed by isolated SNARE proteins is orders of magnitude slower than fusion in living cells^{2,5}.

Cell-free systems that reproduce the events in living cells allow us to dissect

the fusion reaction systematically, thereby placing the SNAREs in a functional context. Because the SNARE proteins involved in

secretion are essential for cell viability, it has been difficult to produce intact membranes that lack one or several of them. One reaction that offers such flexibility is the self, or 'homotypic', membrane fusion of yeast vacuoles. Wickner and colleagues⁶ exploited a simple assay that couples vacuole fusion to activation of alkaline phosphatase, eliciting a diagnostic colour change. The SNARE proteins involved are not essential for cell viability, because homotypic vacuole fusion promotes a dispensable aspect of vacuole growth and inheritance. So, vacuoles have been produced from one yeast strain that lacks a v-SNARE and from another strain that lacks a t-SNARE, to show that fusion requires a v-SNARE in one membrane and a t-SNARE in the other⁶.

Now, Ungermann *et al.*³ and Peters and Mayer⁴ define a set of biochemical requirements that precede, accompany and follow the contact of v- and t-SNARE proteins, a step in the pathway known as docking (Fig. 1). They show that a small GTP-binding protein called Ypt7p tethers vacuoles together in a reversible reaction. Tethering needs to be primed by NSF (*N*-ethylmaleimide-sensitive factor, known as Sec18p in yeast) and α -SNAP (soluble NSF attachment protein, or Sec17p in yeast). But

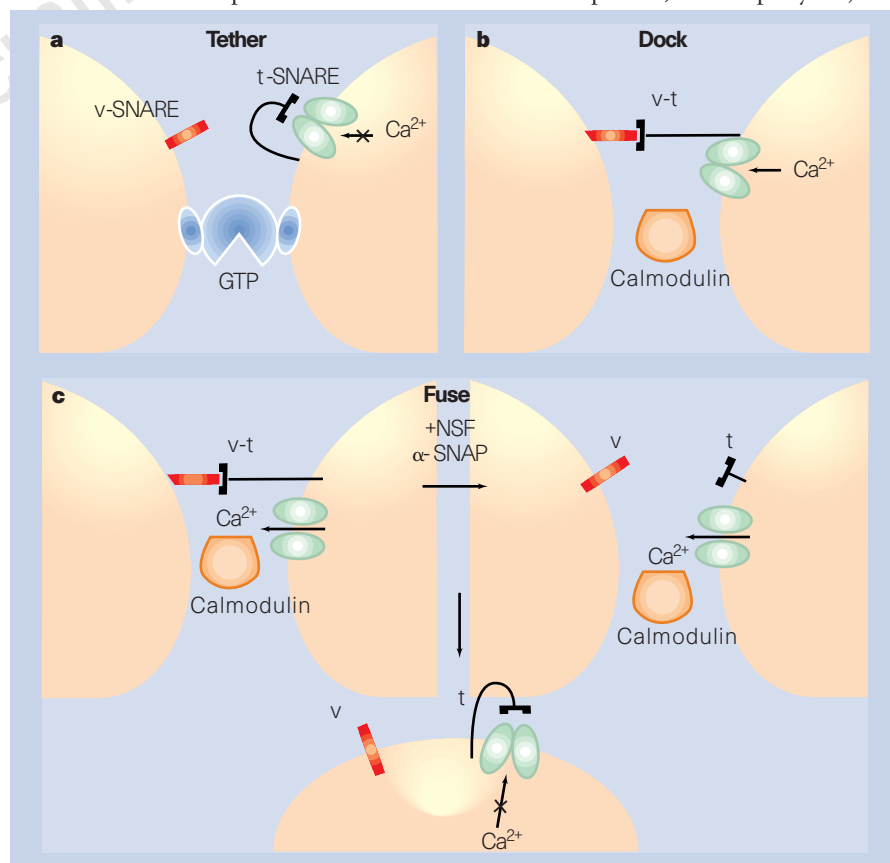


Figure 1 Revised model for membrane fusion, based on the results of Ungermann *et al.*³ and Peters and Mayer⁴. a, A small GTP-binding protein (Ypt7p) tethers together two vacuoles, one bearing a v-SNARE and the other with a t-SNARE. b, Membrane docking triggers a flow of calcium from within the vacuole. c, A calcium–calmodulin complex activates fusion of the vacuoles. Addition of excess *N*-ethylmaleimide-sensitive factor (NSF) and soluble NSF attachment protein (α -SNAP) leads to uncoupling of the SNAREs.

it does not depend on SNAREs. Although we do not know what proteins promote tethering, they must have a loose hold because when yeast Sec19p is added — a protein that binds the GDP-bound form of Ypt7p — the tethered state decomposes. Tethering may be a proof-reading step that allows fusion partners an opportunity to sample and, possibly, to reverse inappropriate encounters. Similar SNARE-independent tethering was first documented for the attachment of transport vesicles derived from the endoplasmic reticulum to the Golgi complex⁷.

At the next stage, membrane docking triggers the flow of Ca^{2+} from a reservoir within the vacuole. Addition of a chelating agent that rapidly sequesters free Ca^{2+} inhibits fusion. Hence, vacuole fusion, like synaptic transmission, is regulated by the local activation of a Ca^{2+} channel. However, unlike at the synapse, a Ca^{2+} -calmodulin complex activates fusion of vacuoles. Although not known, the target of Ca^{2+} -calmodulin must be either the v-t-SNARE complex, or some downstream element of the fusion machine.

Does the v-t complex catalyse fusion or does it activate an as-yet-unknown fusogen? Ungermann *et al.*³ present a provocative experiment consistent with the second view. They show kinetically that the formation of a v-t-SNARE complex precedes, and is separable from, fusion. When they added excess NSF and α -SNAP at the point where v-t-SNARE pairing was maximal, the SNAREs uncoupled with no apparent delay in subsequent membrane fusion. Although the v-t-SNARE complex is an obligatory intermediate in the reaction, these authors conclude that pairing may trigger an independent event that executes fusion. Work by Tahara *et al.*⁸ has shown that SNARE pairs formed on docking of sea-urchin egg secretory vesicles are disassembled by Ca^{2+} at a stage prior to membrane fusion. Of course, in both

cases, the SNARE pair may simply be a non-rate-limiting component of the fusion machine. In any case, the identity of a more direct catalyst for fusion remains to be established.

One obvious suggestion is that docking itself potentiates Ca^{2+} flux. The t-SNARE syntaxin 1A binds and stabilizes the inactivated state of N- and Q-type Ca^{2+} channels in the mammalian central nervous system^{9,10}. Synaptic vesicles may be tethered to the presynaptic membrane until the formation of a v-t-SNARE pair relaxes the inactivated Ca^{2+} channel. The docked complex would then respond to a depolarizing signal and allow a localized Ca^{2+} flux. But Ca^{2+} channels bound by syntaxin and not engaged in a docked state would remain refractory to a depolarizing signal.

Independent evidence indicates¹¹ that syntaxin regulates the cystic fibrosis transmembrane conductance regulator channel. The extreme view would therefore be that SNAREs act mainly by regulating the opening of a channel, and are only indirectly involved in fusion. Genetic and biochemical studies on the functional consequences of the interaction between the vacuole SNAREs and Ca^{2+} channel should allow us to distinguish these possibilities. □

Randy Schekman is in the Department of Molecular and Cell Biology, University of California, Berkeley, California 94720-3202, USA.
e-mail: schekman@uclink4.berkeley.edu

1. Rothman, J. E. *Nature* **372**, 55–63 (1974).
2. Weber, T. *et al.* *Cell* **92**, 759–772 (1998).
3. Ungermann, C. *et al.* *Nature* **396**, 543–548 (1998).
4. Peters, C. & Mayer, A. *Nature* **396**, 575–580 (1998).
5. Nichols, B. J. & Pelham, H. R. B. *Biophys. Biochim. Acta* **1404**, 9–31 (1998).
6. Nichols, B. J. *et al.* *Nature* **387**, 199–202 (1997).
7. Cao, X. *et al.* *EMBO J.* **17**, 2156–2165 (1998).
8. Tahara, M. *et al.* *J. Biol. Chem.* **273**, 667–673 (1998).
9. Sheng, Z. H. *et al.* *Neuron* **13**, 1303–1313 (1994).
10. Bezprozvanny, I. *et al.* *Nature* **378**, 623–626 (1995).
11. Naren, A. P. *et al.* *Nature* **390**, 302–305 (1997).

Volcanology

Remote prospects

William I. Rose and Gregg J. S. Bluth

Robust remote sensing of volcanic emissions has been a long-term goal of Earth scientists interested in the relationship of the solid Earth to the atmosphere, and of volcanologists interested in the telegenic 'reading' of gas emissions to forecast volcanic activity. In the studies that appear on pages 563 and 567 of this issue, Love *et al.*¹ and Francis *et al.*² have taken major steps towards realizing that goal by demonstrating advances in distance (2–10 km) sensing of gas ratios in volcanic plumes. Their work involved the deployment of Fourier-transform infrared spectrometer (FTIR) instruments near Popocatepetl,

Mexico (Fig. 1), and Etna, Italy — two ominously threatening, open-vent volcanoes.

The FTIR instruments take advantage of the spatial frequency characteristics of a spectral image (for example high versus low frequency) to differentiate otherwise obscure features such as minor gas species. The two papers show the feasibility of passive remote sensing using FTIR in the field, aided by solar occultation or using solar radiation scattered by clouds (a key is that solar energy needs to be concentrated in order to detect low burdens of volcanic gas in a high atmospheric background). They also hint at the potential to improve scientific insight into

how volcanoes work; and they will re-ignite interest in field measurements at active volcanoes, which will in turn create many new data on volcanic degassing.

Geochemists have tried extremely hard (and have run great risks) in attempting to directly sample high-temperature, gas-emitting vents on or near volcanoes³. Vents that allow scientific study are rare gifts, however, for the magmatic gases are usually seriously contaminated by chemical constituents of the atmosphere, groundwater and other sources. They usually do not represent a pre-eruptive magmatic gas composition very well, because they are also already substantially degassed. Moreover, all too often it is highly dangerous to visit vents and sample them directly⁴.

Hence the need for remote sensing of volcanic gases. In the past, landmarks have come from the successful ground-based sensing of SO_2 fluxes in volcanic plumes using an independent remote sensor, an ultraviolet correlation spectrometer (COSPEC)⁵, and from satellite spectrometry of the burden of SO_2 in drifting volcanic clouds from large eruptions⁶. Overall, remote sensing of SO_2 has told us a great deal⁷, but it remains of limited value owing to the lack of data on other gas components such as chlorine and fluorine compounds.

Volcanologists have combined the flux measurements with direct sampling where possible^{8,9}, and have also used some sensing methods for other gases that involve entering and profiling volcanic plumes with aircraft¹⁰. Again, however, these techniques cannot be widely used for safety reasons, and adequate sensors to sample gas vents continuously are not yet available.

The new FTIR methods, combined with



Figure 1 The ominously threatening Popocatepetl, which has been active in recent weeks.

NATURE

100 YEARS AGO

While we were taking sympathetic breaths with the insatiable shag, the latter reappeared – yet again with a 15-inch eel. Four 15-inch eels – all swallowed alive – within the space of about four minutes! ... Would he bring up another? Yes, there he was again with another 15-inch eel! A very vigorous eel – just like the others in size and appearance, and swallowed in the same manner, after about 30 seconds' resistance. This made five eels. The question now arose as to what would be the end of this bird. Was he going to die the death of King Henry I before our eyes? ... To make a long story short, we counted *twelve* eels! – all stout 15-inchers. The twelfth seemed, perhaps, rather feeble than the others, but it nearly got away. H.R.H. now seemed to reflect that this last misadventure was a warning, swallowed his twelfth, and took flight... There is, of course, only one explanation of all this; the twelve eels were one and the same eel ... The peculiar procedure of ejecting the prey under water appears very remarkable. From *Nature* 8 December 1898.

50 YEARS AGO

Since last August there have been reports in the Press of a crisis among biologists in the U.S.S.R. The crisis culminated in a decree from the Præsidium of the Academy of Sciences ... This decree dismisses a number of prominent biologists from their posts, closes two famous laboratories, and removes orthodox geneticists from committees... The Lenin Academy of Agriculture Sciences, of which body Lysenko has been president for ten years, held a Conference during July 31–August 7 of this year... Lysenko's address followed familiar lines ... It denies the chromosome theory of heredity. It arraigns several Soviet biologists because their work is inconsistent with Marxist ideology and is sterile of practical results ... The objective of Lysenko's attack on Mendelism is not so much its false ideology, but rather its impotence. By insisting on the stability of the germ plasm Mendelism "condemns practical workers to fruitless waiting". The inheritance of acquired characters is a doctrine necessary for the progress of Soviet agriculture. From *Nature* 11 December 1948.

the SO_2 flux measurements by COSPEC, allow for remote determination of HCl, HF and SiF_4 , and of changes in the flux of several gas species with excellent time resolution. (Incidentally, the authors^{1,2} use different reporting units, which can be related by $\text{tons d}^{-1} \times 0.012 = \text{kg s}^{-1}$).

Application of these new methods will add much to our knowledge of volcano–atmosphere interactions by giving us improved data on many species, particularly the halogen compounds. For example, the study of the ratios $\text{HCl}:\text{SO}_2$ and $\text{HCl}:\text{HF}$ can provide information about the rise and degassing of magma, and is thus a means of eruption monitoring and prediction². Likewise, quantitative measurements of these species can be used to constrain geochemical models of magma equilibrium and oxidation state, and of the interactions among the solid, liquid and gas phases³.

These advances^{1,2} will also challenge our understanding of volcanic gas measurements. In the case of SO_2 , geochemists have had great difficulty in interpreting gas flux data in a consistent way, and many volcanoes have been reported to release 'excess' sulphur^{11,12}; that is, they release more sulphur than expected from study of the mother rock. This uncertainty can apparently now be clarified by a better appreciation of how the oxygen fugacity (effective molecular pressure) in melts influences the separation of SO_2 (ref. 13). An understanding of the pattern or

'order of release' of volcanic gases from magmas as related to their relative solubility has been a long-term guiding idea of work in this area^{14,15} — there is now hope that intrusion of fresh, gas-rich magma into magma bodies, suggested by some interpretations of volcanic gas data¹⁶, might be unambiguously detected by remote sensing and used to increase the accuracy of forecasts. □

William I. Rose and Gregg J. S. Bluth are in the Department of Geological Engineering and Sciences, Michigan Technological University, Houghton, Michigan 49931, USA.

e-mails: raman@mtu.edu
gbluth@mtu.edu

1. Love, S. P., Goff, F., Counce, D., Siebe, C. & Delgado, H. *Nature* **396**, 563–567 (1998).
2. Francis, P., Burton, M. R. & Oppenheimer, C. *Nature* **396**, 567–570 (1998).
3. Symonds, R. B., Rose, W. I., Bluth, G. J. S. & Gerlach, T. Rev. Mineral. **30**, 1–66 (1994).
4. Baxter, P. J. & Gresham, A. *J. Volcanol. Geotherm. Res.* **77**, 325–338 (1997).
5. Stoiber, R. E. & Jepson, A. *Science* **182**, 577–578 (1973).
6. Krueger, A. J. *Science* **220**, 1377–1379 (1983).
7. Bluth, G. J. S. *et al. Nature* **366**, 327–329 (1993).
8. Rose, W. I. *et al. Bull. Volcanol.* **48**, 181–188 (1986).
9. Goff, F. *et al. Geol. Soc. Am. Bull.* **110**, 695–710 (1998).
10. Gerlach, T. M. *et al. J. Geophys. Res.* **102**, 8005–8019 (1997).
11. Andres, R. J. *et al. J. Volcanol. Geotherm. Res.* **46**, 323–329 (1991).
12. Gerlach, T. M. *et al. in Fire and Mud* (eds Newhall, C. G. & Punongbayan, R. S.) 415–434 (Univ. Washington Press, Seattle, 1995).
13. Scaillet, B. *et al. J. Geophys. Res.* **103**, 23937–23949 (1998).
14. Greenland, L. P. *et al. Geochim. Cosmochim. Acta* **49**, 125–129 (1985).
15. Gerlach, T. M. & Graeber, E. *J. Nature* **313**, 273–277 (1986).
16. Harris, D. M. & Rose, W. I. *Bull. Volcanol.* **58**, 163–174 (1996).

Molecular evolution

A hydrogen-producing mitochondrion

T. Martin Embley and William Martin

Dark, damp (and sometimes smelly) places harbour some of nature's most curious eukaryotes. The likes of swamps and intestines are swarming with mostly single-celled eukaryotes (protists) that, like all cells, must produce ATP to survive. Yet these places lack enough oxygen to sustain ATP synthesis as it occurs in textbook mitochondria like our own. Some protists possess no mitochondria, surviving from anaerobic fermentation in the cytosol. Others have quite odd mitochondria that harbour anaerobic ATP-producing pathways. On page 527 of this issue, Akhmanova *et al.*¹ report a gem of such an odd mitochondrion in the ciliate protist *Nyctotherus ovalis*.

The ciliate lives in the suffocatingly oxygen-poor confines of cockroach intestines, where it helps the insect to digest cellulose. Instead of consuming oxygen, *Nyctotherus*'s mitochondrion has the bizarre property of excreting hydrogen as a by-product of ATP synthesis. Similar hydrogen-generating organelles — hydrogenosomes — have been

studied in anaerobic eukaryotes for 25 years². Hydrogenosomes have often been suspected of stemming from the same endosymbiotic bacterium that gave rise to mitochondria. But now, Akhmanova *et al.* report a hydrogenosome that has its own genome, directly betraying its endosymbiotic past.

Cells of *Nyctotherus* (which do not grow in culture and have to be carefully micromanipulated from cockroach hindguts) contain hydrogenosomes that can be labelled by antibodies against DNA. The authors found that the cell produces a ribosomal RNA which, although not proven by *in situ* hybridization to localize to the organelle, bears all the sequence characteristics expected of ciliate mitochondria. Plus, it may look like a mitochondrion, but this DNA-bearing organelle is unquestionably a hydrogenosome because it produces hydrogen — Akhmanova and colleagues found hydrogen-consuming³ methanogenic endosymbionts inside the cells of *Nyctotherus*. Finally, *Nyctotherus* expresses a nuclear-encoded gene for a

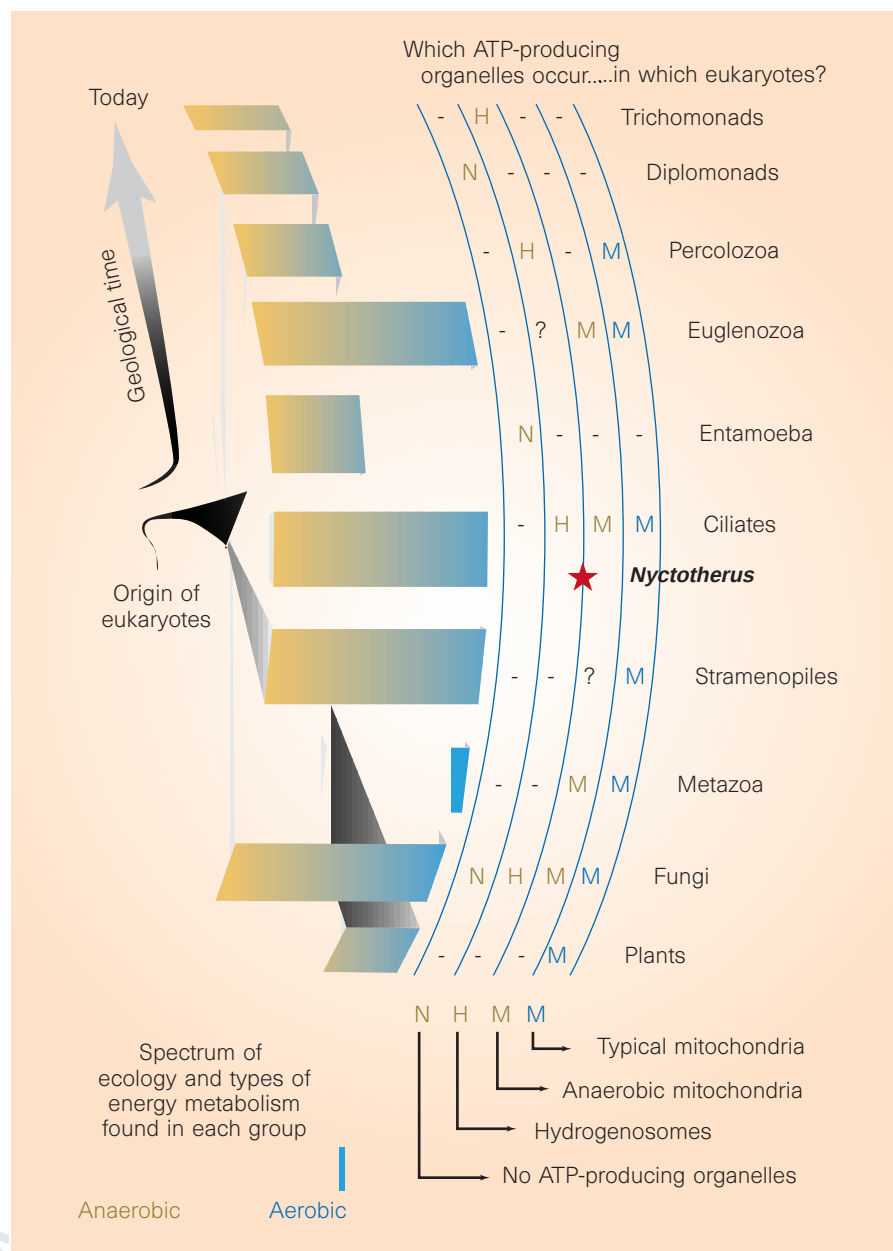


Figure 1 Summary of ATP-producing organelles across a very limited, and arbitrarily chosen, spectrum of eukaryotic taxa. The ciliate *Nyctotherus ovalis* (red star) contains hydrogenosomes, which Akhmanova *et al.*¹ now show contain their own genome. A peculiar group of eukaryotes called microsporidia are related to fungi⁶, whence the 'N' for that group. Antarctic euglenids contain organelles that may be hydrogenosomes¹⁰. Some groups designated 'N' may have organelles, but their role, if any, in energy metabolism is unclear. (Data taken from refs 4, 5 and references therein, and also refs 2, 3, 6–10.)

hydrogenase (an enzyme that makes hydrogen) that is probably imported into the hydrogenosome with a transit peptide¹, as is the case for most proteins in mitochondria.

The evolutionary significance of these findings is twofold. First, hydrogenosome-associated DNA was hitherto a genuine missing link⁴. Such discoveries are rare, and the genes in this DNA are likely to hold exciting surprises. The other, deeper, significance emerges when we remind ourselves that ciliate hydrogenosomes are just the tip of the iceberg (Fig. 1). Hydrogenosomes and the anaerobic (or micro-aerophilic⁵) lifestyle are

widespread among contemporary eukaryotes^{4,6}, including groups that are distantly related in conventional genealogies and, sometimes, arise from within otherwise aerobic groups. Although biologists do not agree which groups of eukaryotes might be the most primitive, most will find their favourite candidate somewhere in Fig. 1, and most biologists do believe that all eukaryotes share a single common ancestor.

The key to understanding hydrogenosomes, and why they produce hydrogen, is energy metabolism⁴. All known eukaryotes generate energy (ATP) by one principle —

Electron acceptor	Excreted end product	Example
O ₂	H ₂ O	Many
Fumarate	Succinate	Flatworm ⁷
NO ₃ ⁻	N ₂ O	Fungus ⁸
NO ₃ ⁻	NO ₂ ⁻	Ciliate ⁹
H ⁺	H ₂	Ciliate ¹

Figure 2 Some electron acceptors, and the corresponding reduced end-products associated with energy metabolism in mitochondria.

the oxidative breakdown of reduced carbon compounds. Electrons removed during the oxidation process must be dumped onto an electron-accepting compound (an acceptor) that can be excreted from the cell. Otherwise, ATP production — and life — comes to a halt. Our mitochondria use oxygen as the acceptor and excrete water. But the mitochondria of anaerobic eukaryotes must resort to compounds other than oxygen (Fig. 2). Some use organic acceptors such as fumarate⁷, some use nitrate^{8,9}. The mitochondria of *Nyctotherus*¹, like other hydrogenosomes^{2,4}, simply transfer the electrons onto protons, producing hydrogen.

Why are anaerobic ATP-producing pathways so widespread in eukaryotes^{2,4,7–10}? There are two popular hypotheses for their origin which, in principle, can be tested through gene-by-gene phylogenetic analysis. One view is that the genes were acquired by eukaryotes through horizontal gene transfer from one or more prokaryotic donors, other than the antecedent of mitochondria. If this were the case, the genes for these pathways in anaerobic eukaryotes should trace to different prokaryotic (eubacterial or archaeobacterial) sources. An alternative view is that the eukaryotic genes involved in anaerobic ATP synthesis were inherited from a single common ancestor of mitochondria and hydrogenosomes. These genes are thought to have been transferred to the host's chromosomes, because they are not found in any known mitochondrial genome. In this case, the common ancestor of contemporary eukaryotes would have acquired, from a facultatively anaerobic protomitochondrion, a genome's worth of genes for all-purpose survival. These genes were then left to the workings of selection and common descent. From this we can predict that, across the anaerobic eukaryotes in Fig. 1, each gene should ultimately trace to a single eubacterial source.

We know that the antecedent of mitochondria brought with it the ability to respire oxygen — much of the respiratory chain is still encoded in mitochondrial DNA¹¹. But we don't know what else was in that symbiont's biochemical repertoire. It

probably existed two billion years ago¹² but, because it's no longer around, we cannot sequence its genome to find out. However, molecular fossils of its lifestyle might be preserved in nuclear chromosomes, including our own, allowing us to piece together the bacterial part of our heritage.

Is anaerobic energy metabolism in eukaryotes a telling relict of our history, or an oddity with little significance? Eukaryotes now regarded as primitive tend to be anaerobes, so these are important questions. The road to answers leads straight to the genomes of eukaryotes from the anoxic world. For scientists studying oxygen-shunning eukaryotes with unconventional mitochondria, with hydrogen-producing mitochondria or with no mitochondria at all, these are exciting times. □

T. Martin Embley is in the Department of Zoology, Natural History Museum, Cromwell Road, London SW7 5BD, UK.

e-mail: tme@nhm.ac.uk

William Martin is at the Institut für Genetik der TU, Spielmannstrasse 7, D-38023 Braunschweig, Germany.

e-mail: w.martin@tu-bs.de

1. Akhmanova, A. *et al.* *Nature* **396**, 527–528 (1998).
2. Müller, M. *J. Gen. Microbiol.* **139**, 2879–2889 (1993).
3. Fenchel, T. & Finlay, B. J. *Ecology and Evolution in Anoxic Worlds* (Oxford Univ. Press, 1995).
4. Müller, M. in *Evolutionary Relationships among Protozoa* (eds Coombs, G. H., Vickermann, K., Sleigh, M. A. & Warren, A.) 109–132 (Chapman & Hall, London, 1998).
5. Lloyd, D. *Nature* **381**, 121 (1996).
6. Embley, T. M. & Hirt, R. P. *Curr. Opin. Genet. Dev.* **8**, 624–629 (1998).
7. Tielens, A. G. M. & Van Hellemond, J. J. *Biochim. Biophys. Acta* **1365**, 71–78 (1998).
8. Kobayashi, M. *et al.* *J. Biol. Chem.* **271**, 16263–16267 (1996).
9. Finlay, B. J., Span, A. S. W. & Harman, J. M. P. *Nature* **303**, 333–336 (1983).
10. Simpson, A. G. B., van den Hoff, J., Bernard, C., Burton, H. R. & Patterson, D. J. *Arch. Protistenkd.* **147**, 213–225 (1997).
11. Gray, M. *et al.* *Nucleic Acids Res.* **26**, 865–878 (1998).
12. Doolittle, W. F. *Proc. Natl Acad. Sci. USA* **94**, 12751–12753 (1997).

Nonlinear physics

Universality of turbulence

Victor S. L'vov

Elsewhere in this issue (*Nature* **396**, 552–554; 1998) Bramwell *et al.* describe the discovery of a new type of universality in turbulence. It connects this strongly non-equilibrium phenomenon of classical fluid mechanics with critical phenomena in the thermodynamic equilibrium of solids (magnetically ordered crystals).

Intuitively, hydrodynamic turbulence is understood as the chaotic motion of fluids — be it of interstellar dust in spiral galaxies, of gaseous planetary atmospheres or of water flowing from a tap (Fig. 1). The length-scales vary from galactic distances of 10^{16} – 10^{18} km, through planetary distances of 1,000–10,000 km down to the human scales of 1 mm–10 m (in the atmosphere and rivers, as well as in the kitchen sink).

Euler's basic mathematical description of fluid dynamics (1741) was corrected to account for viscous friction by Navier (1827)

and Stokes (1945). The Navier–Stokes equation for the velocity $\mathbf{u}(\mathbf{r}, t)$ of fluid at point $\mathbf{r}$ and time t is simply Newton's second law for the fluid particle:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \nu \nabla^2 \mathbf{u} \quad (1)$$

This equates a particle acceleration (the left-hand side) with the forcing due to the gradient of the pressure $p(\mathbf{r}, t)$ and to the viscous friction (the term proportional to the kinematic viscosity of a fluid ν).

In principle, one has to solve this equation to fully understand all turbulent phenomena, but it is a mathematical nightmare. If one ignores the nasty nonlinear term, $(\mathbf{u} \cdot \nabla) \mathbf{u}$, the mean velocity of typical rivers turns out to be about 10^6 km hr⁻¹, and a maximum car velocity is found to be 2,000 km hr⁻¹, both of which are clearly nonsense. The reason is that the nonlinear term is usu-

ally much larger than the linear one. Their ratio is the Reynolds number Re and, for large Re , equation (1) is impossible to solve. Moreover, no one in their right mind wants the full solution of the turbulent velocity field at all points in space-time. It is the statistical properties of the flow, such as probability distribution functions of velocity or the rate of energy consumption, that are important. So, what can we do to understand turbulence?

In 1922, looking at the evolution of turbulent atmospheric conditions, Lewis Fry Richardson suggested the 'cascade picture of turbulence'. In this, the largest eddies in a system are created by instabilities of the mean streamline flow, as in hurricanes. These decay giving rise to eddies of roughly half their size which decay in turn, creating even smaller third-generation eddies — and so on, until the smallest stable eddies lose their energy because of viscous friction that turns it into heat. In high- Re turbulence, eddies exist at various scales, from the largest ones at the scale of the system size down to the smallest ones at the viscous scale. This situation is called 'developed turbulence'.

In 1941 Andrei Kolmogorov estimated the energy $E(R)$ of eddies of scale R in a unit volume of developed turbulence to be $\rho(\bar{\epsilon}R)^{2/3}$. The assumption here is that the only relevant parameter (besides the obvious length scale R and density ρ) is a rate of energy consumption $\bar{\epsilon}$. Kolmogorov's crucial idea was the assumption of universality of small-scale motions (on scales $R \ll L$) in developed turbulence. Here, universality means an independence of the statistical properties of small eddies from the nature of the fluid (be it interstellar dust or water), independence from the mechanism stirring the flow and independence from the particular geometrical form of the container.

But what about the statistics of large-scale motions? Most of us used to believe these were non-universal and would depend on factors such as the geometry of the system. As Bramwell *et al.* now show, however, there is a class of turbulent flows for which at least some characteristics of large-scale turbulent statistics are universal.

The authors demonstrate the rate of power consumption $P(t)$ in turbulent flow in an enclosed air gap between two counter-rotating disks. For very large Re the probability distribution function Q_p is Re -independent when properly rescaled. Their Fig. 1a on page 553 shows plots of $\sigma_p Q_p$ versus $(P - \bar{P})/\sigma_p$ where $\bar{P}$ is the mean value of P and σ_p is the standard deviation of P (a characteristic width of the distribution Q_p). The normalized distributions $\sigma_p Q_p$ for different Re collapse onto the same curve even though they strongly deviate from Gaussian form for $P < \bar{P}$. Bramwell *et al.* argue that just two factors are important: that the integral Re is fixed at a constant value and that flow is

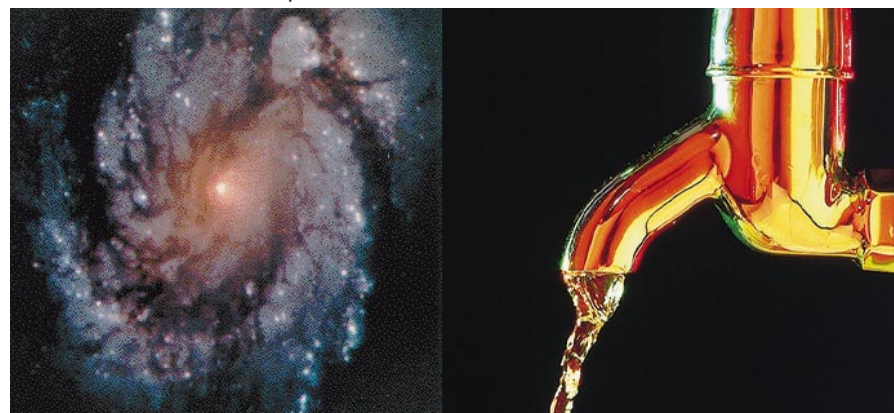


Figure 1 From interstellar space to the kitchen sink. The extreme scales of turbulence range from the motion of interstellar dust in the M100 spiral galaxy to water flowing from a tap.

confined in a closed volume (a cylindrical vessel).

This universality is just the beginning of the story presented by Bramwell *et al.* They also consider a two-dimensional magnetic system at a critical point near the second-order phase transition between magnetically ordered and disordered phases. In such systems the magnetic moment $m(\mathbf{r}, t)$ strongly fluctuates, on a scale with the same probability distribution as velocity fluctuations in developed turbulence. The fluctuations of the total moment $M(t)$ may also be characterized by the probability distribution function Q_M and the authors observe that the functions Q_M and Q_p are amazingly close (see their Fig. 1c on page 553).

Magnetic systems near the second-order phase transition and developed turbulence are microscopically very different physical systems. Moreover, the first one enjoys thermodynamic equilibrium; the second, energy-flux equilibrium. Most surprisingly, however, Bramwell *et al.* show that there are similarities in statistical behaviour between these systems. These results may well serve as a starting point for further theoretical and experimental work to understand the underlying reasons behind such similarities, and their consequences. □

Victor S. Lvov is in the Department of Chemical Physics, Weizmann Institute of Science, Rehovot 76100, Israel.

e-mail: victor@lvov.weizmann.ac.il

Ocean geochemistry

Carbon dioxide uptake at sea

Alain Poisson

To what extent will the world's oceans be able to take up carbon dioxide and reduce its rate of increase in the atmosphere? Numerical modellers^{1,2} have tried to give a provisional answer. But their estimates of the main parameters differ considerably — particularly in the Southern Ocean and the tropics, where the flux of anthropogenic CO₂ across the air–sea interface seems to be largest. On page 560 of this issue³, however, Peng *et al.* claim that it is now possible to make a direct and reliable determination of anthropogenic CO₂ penetration into the ocean. Their approach is based on the temporal variation of measurements of the total dissolved inorganic carbon concentration in seawater.

Before the beginning of the industrial era (around 1850), the global CO₂ cycle was in steady state on a decadal timescale. Since then, the atmospheric concentration of CO₂ has risen increasingly rapidly; now, about

one-half of the CO₂ emitted by the combustion of fossil carbon remains in the atmosphere, the rest being partly dissolved in the oceans and partly stored by the terrestrial biosphere. Researchers have long tried to measure the flux of anthropogenic CO₂ at the air–sea interface and through the oceans' water column, to determine the geographical distribution and rate of CO₂ penetration into deep waters. This is no easy task — data are scarce on the time scales required, and the annual increase of the signal is small relative to the local temporal variability in the oceans' dissolved inorganic carbon.

Twenty years ago, a direct method^{4,5} was devised for estimating the total anthropogenic CO₂ inventory in the ocean. It was based on a simple concept. The total concentration of dissolved inorganic carbon in oceanic deepwater is the result of its original carbon content when it was at the surface, and the changes that the water has subse-

quently undergone. When a parcel of seawater sinks from the surface, its salinity and temperature alter because it becomes mixed with seawater from other origins, implying alteration also of its content of dissolved carbon. Moreover, phytoplankton and zooplankton in the surface layers create organic matter and carbonate skeletons that sink in the water column, are oxidized or dissolved, and produce more dissolved carbon in deep seawater. These changes are respectively estimated by the difference in oxygen content and alkalinity between surface and deep water, and are subtracted from the total inorganic carbon of deep water to obtain the original carbon content of that water when it was at the surface. It was claimed that comparing this original, 'preformed' carbon content with that of the present carbon content of surface water gives the anthropogenic CO₂ signal. But this is only true if the deep water was at the surface in preindustrial times; if it was at the surface after that, only the increase in the signal can be estimated.

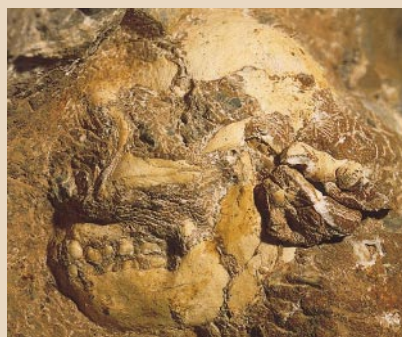
This method was criticized for several reasons, mainly the way of estimating the composition of the preindustrial waters which mix together to form the deep waters, and that of correcting for the changes that occur as they sink. Nevertheless, the profiles of the anthropogenic signal obtained by this simple method are more or less similar to those of chlorofluorocarbons — CFCs, transient tracers whose penetration into the ocean is close to that of anthropogenic CO₂; this approach was later improved by using an exponential distribution of CFCs (ref. 6).

Only recently, however, has the issue of estimating a reliable CO₂ anthropogenic signal been taken up again, largely due to intensification of the collection of carbon data by the World Ocean Circulation Experiment (WOCE) and the Joint Global Ocean Flux Study (JGOFS). The main difficulties are elimination of the nonlinear effects of mixing, and determination of the preindustrial carbon

Palaeontology

The face of Cinderella

The discovery of a complete skull of the hominid *Australopithecus*, associated with abundant limb bone material, is announced this week. The picture shows the exposed left side of the skull. The jaws, with a complete set of teeth in occlusion, can clearly be seen. As R. J. Clarke of the University of the Witwatersrand Medical School in Johannesburg reports (*South African Journal of Science* 94, 460–463; 1998), the skull is still embedded in the Member 2 breccia in the Silberberg Grotto of the Sterkfontein Caves near Krugersdorp, South Africa, and more finds are likely. The taxonomic status of the fossil has yet to be determined, although it is believed to



be more than three million years old.

In 1994, Clarke found four articulating foot bones of *Australopithecus* in rocks from the same site. Eight more foot and

lower leg bones turned up last year, all from the same individual. This discovery prompted Clarke to send his colleagues Nkwane Molefe and Stephen Motsumi into the Silberberg Grotto — like Prince Charming, looking for the girl whose foot would fit a glass slipper — to search for *in situ* remains that would fit neatly onto fragments already found. Cinderella duly turned up, in the form of two lower legs arranged side by side, as if the individual had been buried face-down in the breccia. Further work produced parts of an upper arm and the skull as illustrated. Clarke speculates that the rest of the skeleton is still buried under breccia. **Henry Gee**

R. J. CLARKE

concentration in the original surface waters that contribute to the composition of deep waters. One technique⁷ is based on a quasi-conservative carbon tracer, which reflects the anthropogenic signal, and the disequilibrium of the CO₂ concentration between the atmosphere and the surface water when the water sinks; it is corrected by using the water age estimated by the transient tracers tritium and helium-3. But this method can be applied only locally to specified water layers, and still involves questionable assumptions on the corrections that have to be applied.

Peng *et al.*³ have taken advantage of the high precision of new measurements of the total concentration of dissolved carbon in seawater, and have developed a method for determining the anthropogenic signal on a similar model of variation of the total concentration of dissolved carbon in the ocean. But instead of directly estimating the global inventory of that signal, they calculate its temporal variation at several depths between time periods 17 years apart. This is a good way to eliminate the principal drawbacks of the inventory methods.

This technique is however only valid if the concentration of natural total dissolved carbon and the changes in deep waters due to biological processes do not vary with time. This should be the case, or nearly so, on the decadal time scale. Nonetheless, significant increase of temperature and salinity, correlated with oxygen decrease, has been described for the period 1981–92 in intermediate water in the Atlantic Ocean⁸. So we

cannot be sure whether such variations will affect the calculations.

That apart, the approach of Peng *et al.* gives only the increase of the anthropogenic signal on a decadal scale at specified layers of seawater in specified parts of the ocean. But the high-quality, worldwide data stemming from the WOCE and JGOFS programmes will provide a good basis for extension of that approach, or for use of another new technique⁹ (one which takes the fundamentals of water-source mixing), to produce a direct and reliable global inventory of CO₂. The results will provide powerful constraints on and validation of models of oceanic carbon which aim to estimate the temporal evolution of ocean uptake of anthropogenic CO₂, an essential component for predicting how global climate will change. □

Alain Poisson is in the Laboratoire de Physique et Chimie Marines, Université Pierre et Marie Curie/CNRS, boîte 134, 4 place Jussieu, 75252 Paris Cedex 05, France.

e-mail: apoisson@ccr.jussieu.fr

1. Tans, P. P., Fung, I. Y. & Takahashi, T. *Science* **247**, 1431–1438 (1990).
2. Ciais, P., Tans, P. P., Trolier, M., White, J. W. C. & Francey, R. J. *Science* **269**, 1017–1018 (1995).
3. Peng, T.-H., Wanninkhof, R., Bullister, J. L., Feely, R. A. & Takahashi, T. *Nature* **396**, 560–563 (1998).
4. Brewer, P. G. *Geophys. Res. Lett.* **5**, 997–1000 (1978).
5. Chen, C.-T. A. & Millero, F. K. *Nature* **277**, 205–206 (1979).
6. Papaud, A. & Poisson, A. *J. Mar. Res.* **44**, 385–402 (1986).
7. Gruber, N., Sarmiento, J. L. & Stocker, T. F. *Glob. Biogeochem. Cycles* **10**, 809–837 (1996).
8. Garcia, H., Cruzado, A., Gordon, L. & Escanero, J. *J. Geophys. Res.* **103**, 2817–2830 (1998).
9. Goyet, C. *et al.* *J. Mar. Res.* **57**, 1–30 (1999).

Evolutionary biology

A plastic genome

Pierre Capy

Views on the evolution of genome structure and function have changed dramatically over the past three decades. Once thought to be rather stable, except for occasional changes in chromosome structure brought about by translocations or inversions, the genome is now known to have a flexibility often referred to as plasticity. At the molecular level, the genome is like a puzzle made up of parts that can move from one position to another and, through exchange, deletion, insertion or amplification, generate new combinations of elements with different functions and expression patterns. Indeed, because of this modular construction, unrelated proteins often include similar functional domains.

A remarkable demonstration of plasticity in the *Drosophila* genome is described by Nurminsky *et al.*¹ on page 572 of this issue. They identify a gene that encodes a sperm-specific axonemal dynein intermediate chain, and show that it has evolved extremely rapidly in the lineage leading to *D.*

melanogaster — at least within the three million years or so since the species diverged from *D. simulans*. The new gene is called *Sdic* (for sperm-specific dynein intermediate chain), and it was created from different parts of two unrelated genes; one, *Annexin X* (*AnnX*), encoding a cell-adhesion protein,

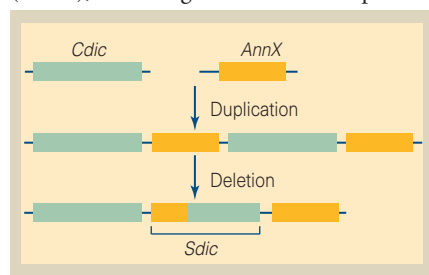


Figure 1 Evolution of the *Sdic* gene in *Drosophila melanogaster*. Nurminsky *et al.*¹ report that *Sdic* evolved recently (within the last three million years) from two unrelated genes. The *Sdic* promoter was created from a protein-coding region of *AnnX*, and its protein-coding exon comes from part of an intron in the *Cdic* gene.

and the other (*Cdic*) encoding a cytoplasmic dynein intermediate chain. In an odd exchange of genetic roles, a protein-coding region of the new gene, and part of an intron in the *Cdic* gene was refashioned into a new protein-coding exon (Fig. 1). The new gene emerged after a duplication and several deletions, and the resulting structure was amplified into a tandemly repeated block of about ten copies.

Sdic is unique among newly evolved genes in *Drosophila* because a putative function has been identified. Another chimaeric gene known as *jingwei* (*jgw*) has been described² in the related species *D. teissieri* and *D. yakuba*. This gene was created in the common ancestor of the two species when a transcript of the gene for alcohol dehydrogenase (*Adh*) was reverse-transcribed into DNA and inserted into the third intron of an unrelated gene called *yande* (*ynd*). Although its function remains unknown, *jgw* is expressed differently in the two species, possibly indicating some divergence in function. Both *Sdic* and *jgw* exemplify 'exon shuffling' in the origin of new genes^{3,4}, but *Sdic* also shows the *de novo* origin of a protein-coding exon from a previously non-coding sequence.

Any new molecular rearrangement will probably be eliminated by deletion unless it is maintained by some sort of selective constraint⁵. Intense selection for modifications in the coding or regulatory regions may lead to one or more 'selective sweeps'. These reduce genetic variability in the new gene itself and, because of 'genetic hitchhiking', also reduce variability around that gene⁶. Indeed, Nurminsky *et al.*¹ analysed the chromosomal region near the newly evolved *Sdic* gene and found an extremely low level of genetic variation. This finding is also consistent with a model of background selection in which the selective elimination of deleterious mutations also eliminates linked neutral variants^{7,8}. Ordinarily there is no way to distinguish between a selective sweep and background selection because the target of the selective sweep is not known⁹. But in the case of *Sdic*, the inference of a selective sweep is reasonably strong in view of independent evidence for large, selectively driven changes in the gene.

Although both *Sdic* and *jgw* were created through genome rearrangements, transposable elements — regions of DNA that can hop around the genome — are not implicated in the origin of either gene. I believe this may be because transposable elements affect genome evolution at a different level. They represent from 5–50% of total genomic DNA, depending on the species, and cause mutations because of their ability to move. Insertion of a transposable element can inactivate a gene, or can change its spatial and temporal expression. Transposition is

estimated to account for about 80% of the mutations detected in *Drosophila* (see ref. 10 for a review). In such insertion mutations, perfect or imperfect excision of the transposable element can completely or partially restore gene function. These elements can also lead to chromosomal rearrangement, such as inversions induced by the *hobo* element in *D. melanogaster*¹⁰, and translocations induced by certain elements in maize¹¹.

Transposable elements are generally considered as selfish DNA, but some can be useful to their hosts¹². As stressed by Ronald Plasterk in these pages¹³, "transposition can occasionally generate gene rearrangements that open new avenues in evolution". On a still longer evolutionary timescale, there may be a connection between introns and transposable elements. For example, transposable elements can be spliced from pre-messenger RNA in maize and *Drosophila*¹⁴. Moreover, analysis of the triosephosphate isomerase genes in insects¹⁵ suggests that some introns may exist owing to insertions of transposable elements. Although studies into the molecular basis of genome evolution are still young, identification of the many processes

in genome evolution, from molecular events to population dynamics, shows the impressive plasticity of the genome and the rapid amplification and fixation of advantageous novelties¹. □

Pierre Capy is in the Laboratoire Populations, Génétique et Evolution, CNRS, 91198 Gif-sur-Yvette Cedex, France.

e-mail: capy@pge.cnrs-gif.fr

1. Nurminsky, D. I., Nurminskaya, M. V., De Aguiar, D. & Hartl, D. L. *Nature* **396**, 572–575 (1998).
2. Long, M. & Langley, C. H. *Science* **260**, 91–95 (1993).
3. Long, M., Souza, S. J. D. & Gilbert, W. *Curr. Opin. Genet. Dev.* **5**, 774–778 (1995).
4. Long, M., Rosenberg, C. & Gilbert, W. *Proc. Natl Acad. Sci. USA* **92**, 12495–12499 (1995).
5. Petrov, D. A., Lozovskaya, E. R. & Hartl, D. L. *Nature* **384**, 346–349 (1996).
6. Maynard Smith, J. & Haig, J. *Genet. Res.* **23**, 23–35 (1974).
7. Hudson, R. R. *Proc. Natl Acad. Sci. USA* **91**, 6815–6818 (1994).
8. Charlesworth, B., Morgan, M. T. & Charlesworth, D. *Genetics* **134**, 1289–1303 (1993).
9. Begun, D. J. & Aquadro, C. F. *Nature* **356**, 519–520 (1992).
10. Capy, P., Bazin, C., Higuier, D. & Langin, T. *Dynamics and Evolution of Transposable Elements* (R. G. Landes, Austin, Texas, 1997).
11. McClintock, B. *Science* **226**, 792–801 (1984).
12. McDonald, J. F. *Trends Ecol. Evol.* **10**, 123–126 (1995).
13. Plasterk, R. *Nature* **394**, 718–719 (1998).
14. Purugganan, M. D. *Trends Ecol. Evol.* **8**, 239–243 (1993).
15. Tyshenko, M. G. & Walker, V. K. *Biochim. Biophys. Acta* **1353**, 131–136 (1997).

Materials science

A precious stone that isn't

Robert W. Cahn

South African jewellers have lately taken to setting a handsome new precious 'stone', a rich purple in colour, in privately commissioned jewellery. The stones (Fig. 1) are made by a straightforward melt-casting approach by Mintek¹, a South African organization specializing in mineral and metallurgical technology. They are in fact not stones at all; rather, they consist of an intermetallic compound, AuAl₂, which is also known to the microelectronic community as the source of the dreaded purple plague.

Plagues and pests have a long history in metallurgy. The preserved food supplies of early expeditions to the polar regions were troubled by tin pest, the gradual conversion

of tetragonal (white) to cubic (grey) tin at low temperatures. Somewhat later, one widespread early method of connecting electrical leads to integrated circuits suffered from the purple plague, a phase that was encountered where gold wires were bonded to aluminium surfaces, or aluminium wires to gold surfaces; it generally led to mechanical failure, either because of the intrinsic brittleness of the phase or because of voids formed during rapid interdiffusion of the two metals^{2,3}. The only solution was to abandon this kind of bond.

A similar problem troubled makers of optical multilayer filters who evaporated gold and aluminium layers in sequence: such multilayers deteriorate rapidly owing to

interdiffusion at room temperature⁴. This diffusion is exceptionally rapid, and is presumably due to the same cause as the rapid age-hardening of aluminium-based aircraft alloys, namely, a high concentration of quenched-in lattice vacancies⁵. Because the atomic radii of gold and aluminium are almost the same, this cannot be a case of 'anomalous rapid diffusion', a process that requires small atoms diffusing among large ones.

The crystal structure of AuAl₂ was determined as long ago as 1934; the phase adopts the simple cubic CaF₂ structure (Fig. 2). Much more recently, Büchler and Range⁶ extracted small single crystals from a mass of the compound which they had synthesized at high pressure and temperature, and found that the Al sublattice is incompletely occupied, so that the actual composition is close to Au₆Al₁₁; the compound contains 'constitutional vacancies'. They also reported that the compound when made in this way is unstable and gradually decomposes at room temperature into a mixture of Al and AuAl; nothing like this has been reported for the 'stones' made by Mintek, which consist of crystal grains some 50 µm across and are said to be "robust and durable" (M. Cortie, Mintek, personal communication). It may be that the instability reported by Büchler and Range is specific to an alloy containing constitutional vacancies, and that these are formed only during high-pressure synthesis.

The intense colour of AuAl₂ has been interpreted in terms of the electronic structure and the allowable interband transitions; the reflectivity falls off in the middle of the visible spectrum but rises again towards the violet end^{7,8}. More recent authors have also interpreted the change of a similar compound, PtAl₂, from yellow to orange-pink as copper is added to the mix; this compound, with its added copper, has also been used a few times for jewellery.

When one looks more closely into the history of AuAl₂, one finds that it has long fascinated metallurgists. It was first reported by W. C. Roberts-Austen (the pioneer after whom 'austenite' in steels was named) in 1892 (ref. 9) and was probably the first intermetallic compound, dental amalgams apart, to have been discovered; he emphasized its remarkable coloration, and went on to point out that its melting point was above that of each of the constituent elements. He claimed that, to the best of his knowledge, AuAl₂ was the only "case of an alloy ... free from mercury, having a higher melting-point than the least fusible of its constituents".

The melting behaviour of alloys in that part of the phase diagram was soon afterwards determined by the Cambridge metallurgists C. T. Heycock and F. H. Neville¹⁰ to be several hundred degrees above that of gold, an indication of its thermodynamic stability. They described AuAl₂ itself as "a magnificent

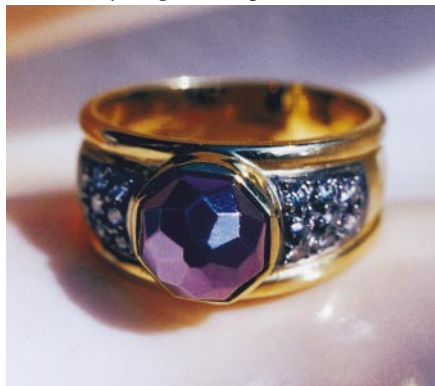


Figure 1 Purple glory — jewellery created by Mintek with the intermetallic compound AuAl₂.

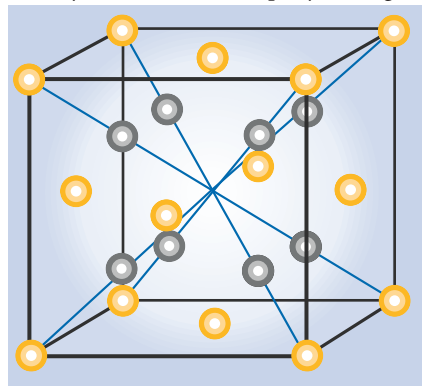


Figure 2 The simple cubic CaF₂ (fluorite) structure, which is also the structure of AuAl₂.

purple body", and elsewhere Neville enthused about the "exquisite eutectic" of AuAl₂ with another Au–Al phase (a eutectic is an intimately mixed microstructure of two distinct crystal types). At the turn of the century, Neville published the first overview¹¹ of intermetallic compounds, almost all discovered in the 1890s; for many, he demonstrated melting temperatures much higher than for the constituent elements.

An intriguing aspect of the use of AuAl₂ and PtAl₂ as precious stones is the fact that they seem to be much less difficult to synthesize than artificial opal, emerald, sapphire and ruby, and they are not used in single-crystal form, unlike those gems; but the new stones differ from those familiar gems in that they each contain a scarce constituent, while the gems are made up of elements such as silicon, oxygen and aluminium.

Perhaps we shall see a growing interest in jewellery incorporating intermetallic com-

pounds containing rare metals. Meantime, although PtAl₂ has been dubbed 'platigem', AuAl₂ used for such purposes does not yet have a common name — why not 'Purple glory'? □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.
e-mail: rwc12@cam.ac.uk

1. www.mintek.co.za
2. Rapson, W. S. & Groenewald, T. *Gold Usage* 293 (Academic, London, 1978).
3. Raymond, T. *Semicond. Int.* **12**, 152 (1989).
4. Hunter, W. R. *et al. Appl. Opt.* **11**, 1594 (1972).
5. Cahn, R. W. & Greer, A. L. in *Physical Metallurgy* (eds Cahn, R. W. & Haasen, P.) 1807–1808 (Elsevier, Amsterdam, 1996).
6. Büchler, H. & Range, K.-J. *J. Less-Common Metals* **160**, 143–152 (1990).
7. Saeger, K. E. & Rodies, J. *Gold Bull.* **10**, 10–14 (1977).
8. Chen, L.-Y. & Lynch, D. W. *Phys. Status Solidi B* **148**, 387–394 (1988).
9. Roberts-Austen, W. C. *Proc. R. Soc. Lond.* **50**, 367–368 (1892).
10. Heycock, C. T. & Neville, F. H. *Proc. R. Soc. Lond.* **66**, 20–21 (1899).
11. Neville, F. H. *Rep. Bradford Mfg Br. Assoc., Sect. B* 1–20 (1900).

Biophysics

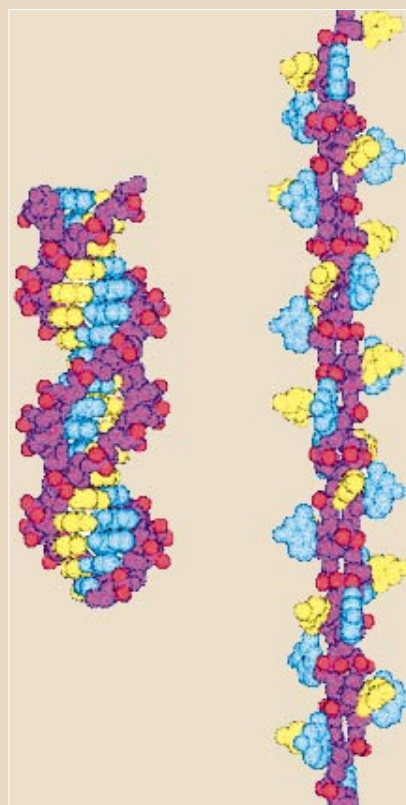
The A to Z of DNA

DNA is remarkably supple — it can be bent, twisted, stretched or squashed into any number of different shapes. The classical, right-handed double helix (pictured right), with 10.4 bases per turn and a phosphate–phosphate distance of roughly 7 Å, is known as B-form DNA. Other stable DNA variants include the A-form structure, in which the distance between phosphate groups is nearer to 5.9 Å, and Z-form DNA, which has a left-handed conformation.

Reporting in *Proceedings of the National Academy of Sciences* (95, 14152–14157; 1998), Jean-François Allemand and colleagues now describe how, through a clever piece of molecular yoga, they have identified another stable DNA conformation. Christened P-form DNA (far right), this new structure is 75% longer than the B form, with just 2.62 bases per turn. Most surprisingly, the phosphate backbone (purple) — which, in B-form DNA wraps around the bases (yellow and blue) — is on the inside.

Allemand *et al.* attached one end of a DNA strand to a flat surface, and the other to a magnetic bead. Then, using magnets to control the position of the bead, they stretched and twisted the DNA. When a force of 3 pN was applied to a positively supercoiled molecule, some of the DNA adopted the new, P-form structure. Such a structure may form *in vivo* too, as the authors believe that these conditions could be re-created during replication and transcription.

A P-DNA-like structure has been found in the genome of the Pfl virus. Moreover, just a few months before Watson and



Crick published their famous description of DNA, Linus Pauling and Robert B. Corey proposed another structure for nucleic acids (*Proc. Natl Acad. Sci.* **39**, 84–97; 1953). Like Allemand *et al.*, Pauling and Corey believed that the "core of the molecule is probably formed of [the] phosphate groups". But although we now know DNA inside-out, it is unlikely that this will be the last new conformation to be discovered.

Alison Mitchell

Daedalus

Vegetable electricity

Electricity seems to be mainly an animal invention. Animals generate it for internal use — crucially in their nerves. Many fish deploy it externally for navigation, and even as a weapon. But plants, which have no nerves, also generate detectable potentials. Tomato plants may even propagate their internal 'wound response' electrically. And while many plants release chemicals which affect others of their species — to prevent local overcrowding, or to 'warn' neighbours of an insect attack — claims have also been made that they can signal electrically.

If so, says Daedalus, the place to look for it is in the sea. It is only the conductivity of the water which allows electric fish to deploy their shocking powers. Seaweeds, like all plants, certainly defend themselves against herbivores. These defences have hitherto been taken as chemical — polyphenols, haloterpenoids and so on, to discourage grazing fish and sea-urchins. But maybe some plants, starting from a fairly feeble electrocommunication ability, have developed it into an effective electrical defence? Intrepid DREADCO botanists with volt-meters and rubber gloves are now trawling the ranks of the seaweeds in search of such powers.

As a first step, they are studying the plant communities which grow on and around undersea electric cables. A weed which senses electricity, or which uses it for terrain-marking or signalling, might well concentrate near such cables. If such a plant is found, the team will scrutinize its related species, hoping to find a seaweed whose voltaic powers are strong enough to stun, or at least discourage, marine herbivores. Even if only mildly electric plants exist in nature, selective breeding could develop serious performers.

An electric plant, of course, would be a revolutionary discovery. It would be the ultimate natural solar energy converter. Small-scale applications would probably develop first. A little tank of water-weed, with each plant connected to a lead and a common counter-electrode in the water itself, might in an isolated community be an ideal self-maintaining 'battery' for a lamp, a radio, a computer, or a mobile telephone. Scale-up to larger 'power ponds' could come later. Daedalus even has visions of glass-bottomed ships propelled by the sunlight shining down through the glass onto its weed coating, and vast lagoons of wired-up seaweed feeding power into the grid.

David Jones

Kendrew constructs; Geis gazes

Today we need computers to create images of the complex structures of protein molecules. Irving Geis, nearly 40 years ago, painted a portrait of myoglobin using only his astounding observational skills.

Martin Kemp

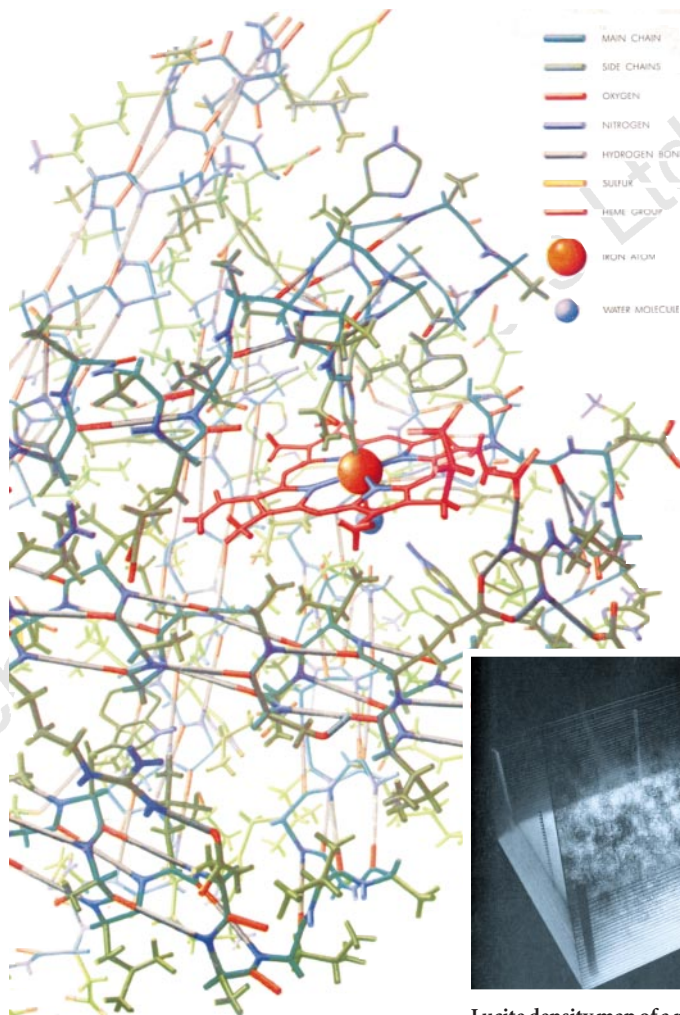
Anyone with even a passing interest in the big molecules that lie at the heart of life knows what they look like on computer screens. Standard imaging packages, such as MolScript, introduced by Per Kraulis in 1991, have made us familiar with the colourful and gently shining tangles of linked atoms, seductively spot-lit and floating weightlessly in space behind the windows of VDUs. What is all too easy to forget is that the basic visual choices in the design of the programs relied on the perceptual and aesthetic visions of the pioneer illustrators during the 1950s and 1960s. Of the illustrators, none was greater than Irving Geis of New York, who died in 1997 at the age of 88.

Having studied architecture, art and design, Geis found his true *métier* as a leading illustrator and artist of science working for *Scientific American* from 1948 onwards. It was in this capacity that he created his first masterpiece of protein portraiture for John Kendrew's article "The Three-dimensional Structure of a Protein Molecule", published in December 1961 (Vol. 205, pages 96–110), which remains a classic of visual and verbal exposition. In the bound annual volume I consulted, Kendrew's article was immediately identifiable by its thumbled pages, dog-eared corners and torn edges reinforced by tape.

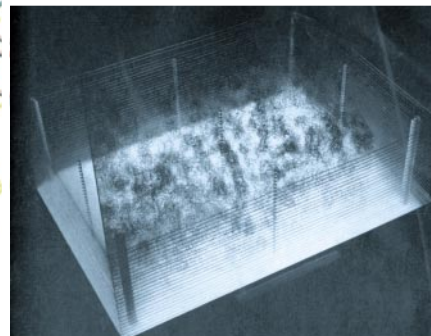
Visual and verbal excitement go perfectly hand in hand. Kendrew begins, "When the early explorers of America made their first landfall", and concludes that "students of the living organism do indeed stand on the threshold of a new world". Kendrew and Geis avail themselves of virtually every kind of available technique to portray their new world.

There are flat diagrams of molecular skeletons, linear maps of electron density, line drawings of ball-and-rod structures, X-ray photographs and photo-micrographs, together with photographs of equipment and of various kinds of models composed of clips and rods, Lucite sheets and a lumpy sculpture. And, above all, there is Geis's remarkable painting.

The subject was the 2,600-atom molecule of sperm whale myoglobin, of which Kendrew had caught his first sight on a Sunday in May 1957. At first he worked to a resolution of six angstroms, but determined that a higher resolution was needed — "as in the case of a musical note,



Detail of the model of the sperm whale myoglobin molecule, painted by Irving Geis.



Lucite density map of a myoglobin crystal, constructed by Kendrew *et al.*

the greater the number of higher harmonics that are included, the sharper and more precise is the resulting picture". The two-angstrom scale he finally used involved two sets of 10,000 X-ray reflections. The refined data were used to build the astonishing three-dimensional contour map of electron density on 50 layered sheets of Lucite — at the cost of six man-months of work.

"At first sight" the density map "seemed completely irregular", but just as "driving past an orchard" will disclose different orders from different perspectives, so Kendrew found that if he "looked through the stack of Lucite sheets in a direction corresponding to the axis of one of the rods" (the concentrations of higher density already observed at lower resolution), the rods could be seen as hollow and spiral in configuration.

The next step was the construction of a model composed of a forest of uprights with coloured clips which represented the

whole molecule with every side chain in place. It was through this forest that Geis gazed with remorseless concentration when painting his portrait of the molecule, using his unrivalled command of perspective, light and shade, and colour recession to reveal the intricate sculptural web of linkages. Looking at it 37 years later, it is difficult to credit that it was not done by computer.

Indeed, there is a case for saying that our current computer graphics, for all their novel features such as animation and stereoscopy, are basically providing a series of powerful, dynamic and convenient glosses on the graphic modes invented by Geis and his fellow pioneers of hand-made representation. □

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.

A hydrogenosome with a genome

Some anaerobic protozoa and chytrid-omycete fungi possess membrane-bound organelles known as hydrogenosomes. Hydrogenosomes are about 1 micrometre in diameter and are so called because they produce molecular hydrogen¹. It has been postulated that hydrogenosomes evolved from mitochondria by the concomitant loss of their respiration and organellar genomes¹⁻⁴, and so far no hydrogenosome has been found that has a genome^{1,2}. Here we provide evidence for the existence of a hydrogenosomal genome of mitochondrial descent, and show that the anaerobic heterotrichous ciliate *Nyctotherus ovalis* possesses a new type of nuclear-encoded 'iron-only' hydrogenase enzyme.

N. ovalis, found in the hindgut of the cockroaches *Periplaneta americana* and *Blaberus* spp.⁵, has numerous hydrogenosomes that are intimately associated with endosymbiotic methane-producing Archaea, which use hydrogen produced by the hydrogenosomes (Fig. 1a). The hydrogenosomes are bounded by distinct double membranes and have an inner membrane with

cristae-like projections. The matrix contains ribosome-like particles of the same size as the numerous 70S ribosomes of the endosymbiotic methanogenic Archaea (Fig. 1d).

Weak but consistent immunogold labelling was obtained with a commercial antiserum against DNA in more than 80% of the hydrogenosomes we sectioned (Fig. 1b). We labelled the same organelles by using heterologous antisera against an iron-only hydrogenase ([Fe]-hydrogenase⁶) and a hydrogenosomal adenylate kinase of the AK2 type (Fig. 1c). Electron microscopy and immunocytochemistry indicated that DNA, ribosomes and components of a hydrogenosomal metabolism are present in the hydrogenosomes of *N. ovalis*.

Because all hydrogenosomes studied so far lack a genome^{1,2}, we wondered whether the immunoreactive organelle DNA in *N. ovalis* is functional. By using the polymerase chain reaction (PCR), with primers directed against conserved regions of the mitochondrial small-subunit (SSU) ribosomal RNA genes from ciliates^{7,8}, we isolated and

cloned a fragment of a homologous SSU rRNA gene from *N. ovalis* total DNA. The complete sequence of this rDNA was obtained by rapid amplification of complementary DNA ends, using total RNA from *N. ovalis* as a template. Phylogenetic analysis of conserved regions places the *N. ovalis* sequence with high bootstrap values within the mitochondrial SSU rRNA genes from aerobic ciliates (not shown). Northern blotting reveals that the isolated SSU rRNA gene is abundantly expressed. The cross-hybridizing rRNA is fragmented, supporting the idea of descent from a mitochondrial rRNA gene of a ciliate^{7,8} (Fig. 2a).

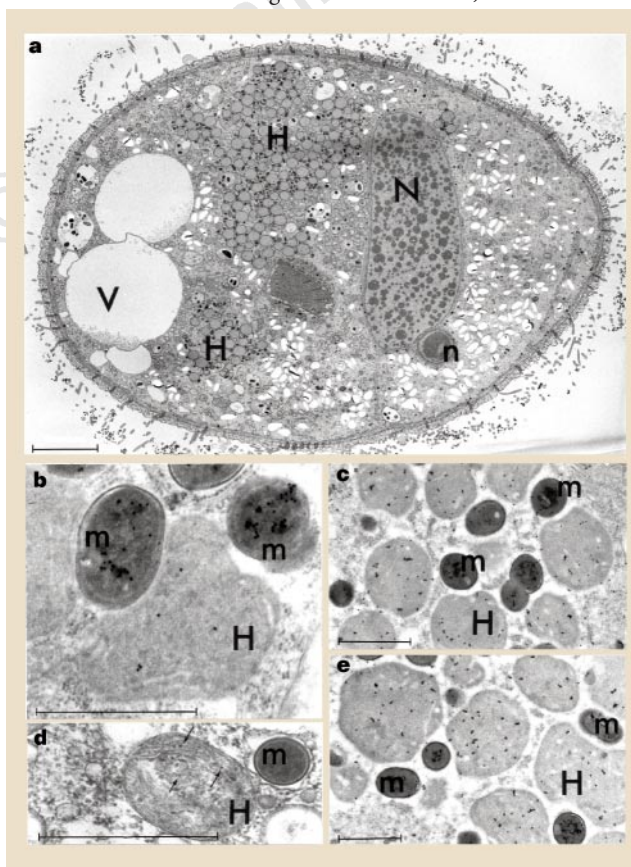
By using PCR, we identified a nuclear-encoded cDNA 3.6 kilobases long encoding a putative [Fe]-hydrogenase. The cDNA contains a single open reading frame (ORF) consisting of 1,198 codons (Fig. 2e). The amino-terminal half of the predicted polypeptide (residues 22–620) shares 35–41% identity with the [Fe]-hydrogenases from the bacterium *Clostridium* and the proteobacterium *Desulfovibrio*⁹. The middle part of the ORF (residues 630–810) shares 21–24% identity with the HoxE protein of *Synechocystis* spp., the NADP-reducing hydrogenase subunit HndA of *Desulfovibrio fructosovorans*, and the (nuclear-encoded) NuoE/Nuo5 precursors of the 24K (relative molecular mass 24,000) subunit of complex I (NADH-ubiquinone oxidoreductase) from the respiratory chain. The carboxy-terminal part of the ORF (residues 840–1,170) has 28–34% identity with the HoxF genes of *Synechocystis* spp. and *Alcaligenes eutrophus*, and, notably, with the nuclear-encoded NuoF/Nuo6 precursors of the 51K subunit of mitochondrial complex I. In all subunits, the NAD-binding, flavin-mononucleotide-binding and Fe-S motifs are conserved.

The hydrogenase cDNA hybridizes to a 4-kilobase fragment of undigested genomic DNA (Fig. 2d), indicating that the hydrogenase is encoded by gene-sized pieces of *N. ovalis* macronuclear DNA¹⁰ (Fig. 2c). The genomic fragment terminates in a G₅T₄G₅(T₄G₄)₅ repeat that is very similar to the telomere sequences of hypotrichous ciliates¹⁰. The cDNA start was about 180 base pairs downstream from the telomere. The 16 amino-terminal amino acids of the ORF resemble mitochondrial transit peptides. The length of the *N. ovalis* hydrogenase mRNA is in agreement with the sequence data (Fig. 2b).

Our results indicate that *N. ovalis* hydrogenosomes evolved from mitochondria but, contrary to recent predictions², they have not relinquished their genome. The evolutionary origin of the chimaeric *N.*

Figure 1 Electron micrographs of *Nyctotherus ovalis* showing hydrogenosomes.

a, *N. ovalis* from the hindgut of *Blaberus*⁵. The hydrogenosomes (H) are surrounded by endosymbiotic methane-producing Archaea (dark spots); N, macronucleus; n, micronucleus; V, vacuole. Visible by MnO₄ fixation/Epon. **b, c, e**, *N. ovalis* from the hindgut of *Periplaneta americana*⁵. Immunogold labelling of glutaraldehyde-fixed and Unicryl-embedded sections; m, methanogenic Archaea (endosymbionts). **b**, DNA antiserum (Boehringer) labels the matrix of about 80% of the hydrogenosomes on randomly chosen sections with 3–10 grains. The difference in DNA concentration causes the label over the endosymbiotic methanogens to be heavier. **c**, Immunogold labelling obtained with a polyclonal antiserum against hydrogenosomal adenylate



kinase (hdgAK2L2) from the anaerobic chytrid *Neocallimastix* sp. L2 (F. V. and B. B., unpublished). Matrix of hydrogenosomes and endosymbiotic methanogens is labelled. **e**, Labelling of the hydrogenosomes (and methanogens) with an antiserum against an [Fe]-hydrogenase of *Trichomonas vaginalis*⁶. There is about 40% amino-acid sequence identity with the [Fe]-hydrogenase described here. **d**, Electron micrograph of a hydrogenosome of *N. ovalis* from *Blaberus*. OsO₄ fixation/Epon. Cristae are clearly visible. Arrows indicate ribosomes. Scale bars are 1 μ m, except for that in **a**, which represents 10 μ m.

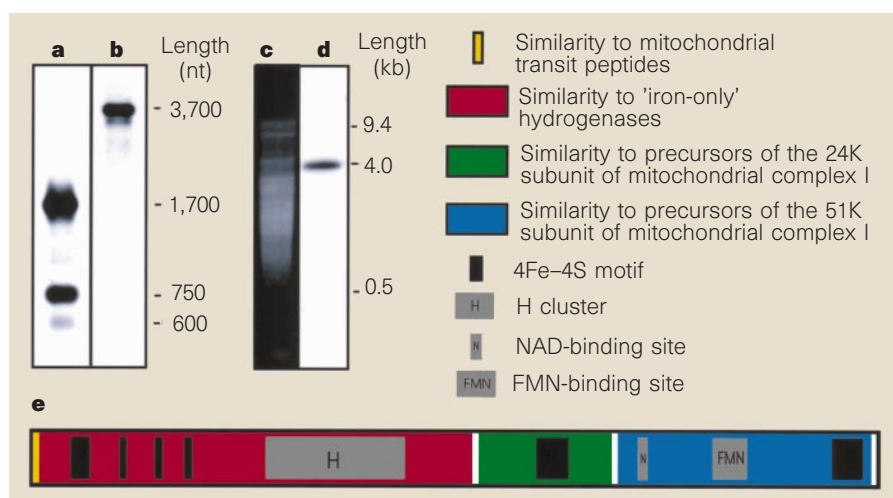


Figure 2 The *N. ovalis* genome. **a**, Northern blot of *N. ovalis* total RNA hybridized to a ^{32}P -labelled fragment of the mitochondrial SSU ribosomal RNA (positions 550–1660). The largest hybridizing RNA species corresponds in size to the sequenced mitochondrial SSU rRNA, at 1,701 nucleotides (nt). The smaller RNA fragments on the blot probably represent naturally occurring discontinuities, similar to those described for *Paramecium* and *Tetrahymena*²⁸. **b**, Northern blot of *N. ovalis* RNA hybridized to the 3' part of the hydrogenase complementary DNA (positions 1,308–3,625). An identical result was obtained when the 5' part of the cDNA was used as a probe. **c**, Ethidium bromide-stained gel with 5 μg undigested genomic *N. ovalis* DNA. Most DNA fragments are gene-sized, at 0.5–10 kilobases (kb). **d**, Southern blot of the same gel hybridized to the ^{32}P -labelled hydrogenase cDNA fragment (positions 1,308–3,625). **e**, The open reading frame encoding the putative [Fe]-hydrogenase. Homologies to known proteins and putative functional domains are indicated. Motifs not to scale.

ovalis hydrogenase gene remains puzzling. Sequence similarity suggests that the hydrogenase couples hydrogen production to the reoxidation of NADH through a combination of functional components derived from respiratory (complex I modules) and fermentative ([Fe]-hydrogenase module) metabolism. The hydrogenase gene could have been inherited from the common proteobacterial ancestor of mitochondria and hydrogenosomes³; it may have been acquired through lateral gene transfer from prokaryotes in the *N. ovalis* lineage; or perhaps it was assembled *de novo* in the *N. ovalis* lineage from pre-existing or acquired genetic components.

Hydrogenase expression is clearly a prerequisite for the conversion from mitochondrion to hydrogenosome during the eukaryotic specialization to anaerobic niches. Because hydrogenosomes have arisen independently several times in mitochondrion-bearing lineages^{1,4,11}, it is possible that their hydrogenases did as well.

Anna Akhmanova*†, Frank Voncken*, Theo van Alen*, Angela van Hoek*, Brigitte Boxma*, Godfried Vogels*, Marten Veenhuis‡, Johannes H.P. Hackstein*

*Department of Microbiology and Evolutionary Biology, University of Nijmegen, NL-6525 ED Nijmegen, The Netherlands
e-mail: hack@sci.kun.nl

‡Department of Microbiology, University of Groningen, NL-9750 AA Haren, The Netherlands.

†Present address: Department of Cell Biology and

Genetics, Erasmus University, NL-3000 DR Rotterdam, The Netherlands

1. Müller, M. J. *Gen. Microbiol.* **139**, 2879–2889 (1993).
2. Palmer, J. D. *Science* **275**, 790–791 (1997).
3. Martin, W. & Müller, M. *Nature* **392**, 37–41 (1998).
4. Fenchel, T. & Finlay, B. J. *Ecology and Evolution in Anoxic Worlds* (Oxford Univ. Press, 1995).
5. van Hoek, A. H. A. M., van Alen, T. A., Sprakel, V. S. I., Hackstein, J. H. P. & Vogels, G. D. *Mol. Biol. Evol.* **15**, 1195–1206 (1998).
6. Bui, E. T. N. & Johnson, P. J. *Mol. Biochem. Parasitol.* **76**, 305–310 (1996).
7. Seilhamer, J. J., Olsen, G. J. & Cummings, D. J. *J. Biol. Chem.* **259**, 5167–5172 (1984).
8. Schnare, M. N., Heinonen, T. Y., Young, P. G. & Gray, M. W. *J. Biol. Chem.* **261**, 5187–5193 (1986).
9. Casolot, L. *et al. Anaerobe* **4**, 45–55 (1998).
10. Hoffman, D. C., Anderson, R. C., Dubois, M. L. & Prescott, D. *M. Nucl. Acids Res.* **23**, 1279–1283 (1995).
11. Embley, T. M. *et al. Proc. R. Soc. Lond. B* **262**, 87–93 (1995).

Sequence data have been deposited in the EMBL database under accession numbers Y16775, Y16669 and Y16670.

Did dinosaurs come up to scratch?

The high specificity of bird ectoparasites has frequently been interpreted to mean that they have a long evolutionary history¹. Either they were present as parasites before the main diversification of birds in the Cretaceous period², or they evolved independently into avian parasites on several occasions. The recent discovery of dinosaurs with feathers³ suggests that birds may have inherited some of their ectoparasites from feathered theropod dinosaurs. We have now found microscopic egg-like structures on the surface of a fossil feather

of approximately 120 million years old from the Crato Formation (Lower Cretaceous period) of northeast Brazil. They closely resemble the eggs of ectoparasitic feather mites and indicate an early origin for this specialized group of Acari.

Modern birds have diverse and host-specific ectoparasites^{4,5}, with many birds being host to several different orders on different parts of the body⁶. The earliest undisputed fossil feathers, of *Archaeopteryx* from the Upper Jurassic period of Bavaria, Germany⁷, indicate that ectoparasites could have been using feathers for more than 140 million years. The recent discovery of maniraptoran dinosaurs bearing feathers³ is suggestive of an even older record, but direct evidence of feather-borne ectoparasites in the Mesozoic era has been lacking.

Here we describe a fossil feather from a famous fossil Lagerstätte, the Nova Olinda Member of the Crato Formation⁸, and provide direct evidence that Mesozoic feathers carried parasites. The specimen (National Science Museum of Japan, Tokyo, specimen NSM PV20059) is a slightly asymmetric, probable caudal, plume 85 millimetres in length and with a maximum width of 11 millimetres (Fig. 1a,b).

On the most well-preserved slab there are more than 100 hollow, sub-spherical structures (Fig. 1c), with diameters of 68 to 75 micrometres, adhering to the surface of the rachis and barbs. They resemble the eggs of parasitic Acari⁹. Some spheres are isolated on a single barb, whereas others appear to be concentrated in loose clusters of several spheres. Where examples have split to allow measurement, the walls were between 5 and 7 micrometres thick. They are reddish brown and slightly shiny, which is similar to the preservation style seen in fossil insects from the same deposits⁸.

We believe that they are eggs rather than artefacts of preservation because they have a consistent morphology and size; they are restricted to the feather; they are an order of magnitude larger than pyrite framboids, a common spheroidal diagenetic product; and, because of their hollow nature and the circular aperture, 35 to 40 micrometres across (seen on a few of the spheres), are inconsistent with a diagenetic origin. We rule out a spore or pollen affinity on the basis of morphology and the fact that spores and pollens (common in this deposit) are usually organic, rather than preserved as limonitic replacements.

Birds are common hosts to several ectoparasite groups, including fleas, mites and ticks². All three of these arthropod groups lay eggs, but fleas do not usually lay their eggs on the host. Bird lice (order Phthiraptera) cement their eggs to the surface of feathers, with some species being specific to certain parts of the feather, such as the shaft or vane^{10,11}. However, the eggs

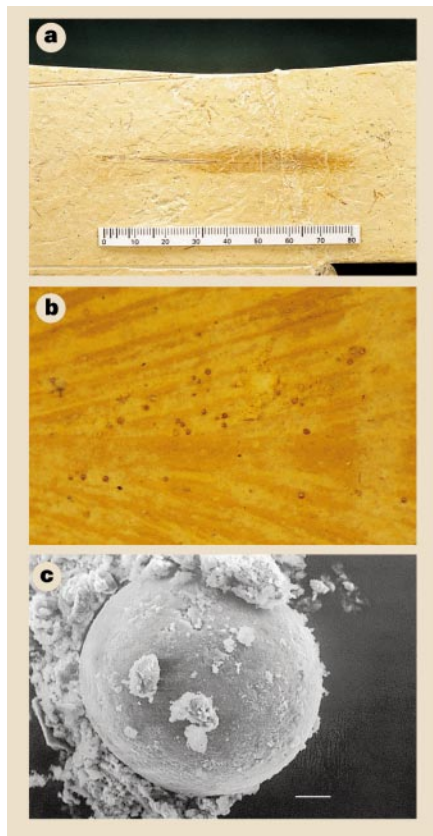


Figure 1 Fossil feather with ectoparasite eggs. **a**, Fossil feather from the Lower Cretaceous (Aptian) Crato Formation, Nova Olindá region, Ceara, Brazil. The feather is 85 mm long. **b**, Detail of the central portion of the feather showing loose clusters of probable ectoparasitic acariid eggs. Individual eggs are between 65 and 70 micrometres in diameter. **c**, Scanning electron micrograph of a probable ectoparasitic acariid egg from the feather. Scale bar, 10 micrometres.

described here are much smaller (70 micrometres in diameter) than the average size for eggs of phthirapterans (more than 500 micrometres), and are more similar to the eggs of small mites (Acari). Furthermore, the eggs are almost spherical, like those of Acari, rather than elongate and barrel shaped, like those of lice. However, several Acari families parasitize birds, and we are unable to identify the family to which the eggs belong.

The oldest record of feather-borne parasites is currently of a mallophagan (feather louse) egg on a feather entombed in Baltic (Oligocene) amber¹². Our specimens indicate that feather-borne parasites have at least a Lower Cretaceous origin and may be even older, with a co-evolutionary history that perhaps spans more than 140 million years.

David M. Martill, Paul G. Davis

School of Earth, Environmental and Physical Sciences, University of Portsmouth, Portsmouth PO1 3QL, UK
e-mail: david.martill@port.ac.uk

1. Mitter, C. & Brooks, D. R. in *Coevolution* (eds Futuyma, D. J. & Slatkin, M.) 65–98 (Sinauer, Sunderland, MA, 1983).

- Boucot, A. J. *Evolutionary Paleobiology of Behavior and Coevolution* (Elsevier, Amsterdam, 1990).
- Ji, Q., Currie, P. J., Norell, M. A. & Ji, S.-N. *Nature* **393**, 753–761 (1998).
- Clayton, D. H. & Moore, J. (eds) *Host-parasite Evolution, General Principles and Avian Models* (Oxford Univ. Press, 1997).
- Janovy, J. Jr in *Host-parasite Evolution, General Principles and Avian Models* (eds Clayton, D. H. & Moore, J.) 303–337 (Oxford Univ. Press, 1997).
- Askew, R. R. *Parasitic Insects* (Elsevier, London, 1971).
- Wellnhofer, P. *Archaeopteryx* **11**, 1–48 (1993).
- Martill, D. M. *Field Guides to Fossils Vol. 5. Fossils of the Santana and Crato Formations, Brazil*. (Palaeontological Association, London, 1993).
- Lapage, G. *Veterinary Parasitology* 2nd edn (Oliver & Boyd, Edinburgh, 1968).
- Atyeo, W. T. & Windingstadt, R. M. *J. Parasitol.* **65**, 650–658 (1979).
- Coles, B. H. *Avian Medicine and Surgery* (Blackwell, Oxford, 1985).
- Bachofen-Echt, A. *Der Bernstein und seine Einschlüsse* (Springer, Berlin, 1949).

Increased summertime heat stress in the US

In the past half century, the mean summertime temperature in the United States has increased^{1–4}, with nights warming more than days^{5,6}. When humidity is high, hot weather can cause heat stress in humans. Here we show that the frequency of extreme heat-stress events in the United States, caused by extremely hot and humid days as well as by heatwaves lasting for several days, has increased over the period from 1949 to 1995.

We expect increases in mean summertime temperature to be accompanied by an increased frequency of days with extremely high temperature (T) (ref. 7), although there have been few attempts⁸ to find observational evidence. Extremes of summertime heat have a greater impact on human health than any other severe weather in the United States⁹, with elderly people being most at risk¹⁰.

The combined effects of T and humidity in sultry weather are incorporated in the 'apparent temperature', A (ref. 11), which is a widely used measure of human heat stress. A can be expressed in terms of ambient T and water-vapour pressure (e , in kilopascals) as $A = -1.3 + 0.92 T + 2.2 e$ (ref. 11). In the United States, mean summertime A is increasing faster than mean summertime T because humidity has increased by several per cent per decade⁴.

Analysis of mortality statistics and weather data for a few dozen cities in the United States has revealed that there are threshold T (ref. 10) and A (J. Detwiler, R. Livezey and L. Kalkstein, unpublished data) values above which mortality increases sharply. These threshold values are closely correlated with the 85th percentile values of T and A derived from climatological data (see Supplementary Information). Here we use climatology-based thresholds of A and T from 113 US weather stations (mainly at airports) to define events of extreme heat stress. The thresholds are locally defined on the basis of three-hourly temperature and humidity data for 1961 to 1990. Separate thresholds were computed for daily-average, daily-maximum and daily-minimum A and T (see Supplementary Information for a map of the thresholds of daily-average A). The thresholds are based on data for July and August (which are the hottest months), but they can be exceeded during any month of the year: they are typically exceeded about 12 times a year.

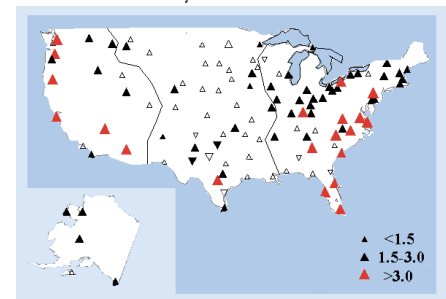


Figure 1 Trends in the annual frequency of the daily-minimum apparent temperature exceeding local threshold values from 1949 to 1995. The size and colour of the triangle indicates the magnitude of the trend (ranging from -2.7 to $+5.2$ per decade); its orientation (on its base or its apex) indicates the sign of the trend (positive or negative, respectively). Filled triangles indicate significant trends (at $P < 0.05$, using non-parametric methods)⁴⁴.

Table 1 Trends in extreme heat events in the United States

	United States	Regional decadal trends		
		Western	Central	Eastern
Days with $A_{ave} > A^*$	1.67	2.08	1.19	2.06
Days with $A_{max} > A^*$	0.61	1.13	0.46	0.56
Days with $A_{min} > A^*$	1.94	2.34	0.79	2.60
Three-day heatwaves with $A_{ave} > A^*$	0.32	0.37	0.22	0.39
Four-day heatwaves with $A_{ave} > A^*$	0.23	0.27	0.15	0.28
Days with $T_{ave} > T^*$	1.56	2.14	0.92	1.81
Days with $T_{max} > T^*$	0.29	1.13	0.20	0.03
Days with $T_{min} > T^*$	1.79	2.41	0.41	2.48

Decadal trends in the annual frequency of extreme apparent temperature, A , and temperature, T , for the United States as a whole and for three regions of the country from 1949 to 1995. Trends significant for $P < 0.05$ or less are shown in bold. Extreme events are based on local thresholds, defined by the 85th percentile values of July and August daily averages (A^* and T^*), daily minima (A_{min}^* and T_{min}^*) and daily maxima (A_{max}^* and T_{max}^*). These climatologically based thresholds are closely correlated with thresholds associated with increased mortality (see Supplementary Information).

From 1949 to 1995, the annual frequency of days exceeding the thresholds increased at most stations. These increases are largest for daily-minimum *A*, and the number of high heat-stress nights increased by 25% or more at some locations. The largest and most statistically significant trends occurred in some of the most populated areas, in the eastern and western thirds of the United States (Fig. 1).

The spatial distribution of the trends allows us to present results for three regions of the country (Fig. 1) bounded approximately by the Mississippi River and the continental divide (Rocky Mountains). Regional trends in the frequency of extreme daily-maximum *A* are substantially smaller than those for extreme daily minima, and are statistically significant only in the western region (Table 1). In the eastern and central United States, trends in extreme *A* are larger than trends in extreme *T*. In the western region, where summertime humidity increases are less marked⁴, the *A* and *T* trends are similar.

These increases in single-day heat stress events are associated with increases in heatwaves, defined as runs of three or four consecutive days with daily-average *A* exceeding the 85th percentile value. On average, each weather station experiences 1.7 three-day and one four-day heatwaves per year. Upward trends in the frequency of heatwaves are highly significant ($P < 0.01$) in the eastern and western regions (Table 1) and indicate an increase of about 20% in the number of heatwaves over the period from 1949 to 1995.

These trends may be partly associated with increased urbanization. If the spatial extent of urban heat islands has been growing, weather stations at airports near large cities might experience high temperatures more frequently, especially at night¹². However, the regional consistency of the trends suggests that their origins are not strictly local.

If these climate trends continue they may pose a public health problem¹³, particularly as there are increasing numbers of elderly people, who are most vulnerable to heat-related sickness and mortality¹⁰.

Dian J. Gaffen, Rebecca J. Ross

National Oceanic and Atmospheric Administration,
Air Resources Laboratory,
1315 East West Highway,
Silver Spring, Maryland 20910, USA

(1989).

11. Steadman, R. G. *J. Clim. Appl. Meteorol.* **23**, 1674–1687 (1984).
12. Gallo, K. P., Owen, T. W., Easterling, D. R. & Jameson, P. F. *J. Clim.* (in the press).
13. McMichael, A. J. *Climate Change 1995: Impacts, Adaptations and Mitigation of Climate Change* (Cambridge Univ. Press, 1996).
14. Lanzante, J. R. *Int. J. Climatol.* **16**, 1197–1226 (1996).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

p53 polymorphism and risk of cervical cancer

Storey and co-workers¹ recently presented results indicating that the allele encoding arginine in the codon-72 polymorphism of the *p53* gene represents a significant risk factor in the development of cancers associated with human papilloma virus (HPV). The form of the *p53* protein carrying an arginine residue at this position was found to be significantly more susceptible to degradation by the HPV E6 protein than by the proline form. Genotype analysis of 30 cervical tumours and 12 skin carcinomas revealed that the homozygous Arg/Arg genotype was overrepresented compared with 41 controls. We have now analysed this polymorphism in leukocyte DNA from a larger sample of cancer patients and controls but have found no significant overrepresentation of this genotype.

We analysed leukocyte DNA from 77 cervical-cancer patients² who were positive for 'high-risk' HPVs and 92 patients with cervical intraepithelial neoplasia (CIN) grades II–III, of which 72 were positive for 'high-risk' HPVs. For controls, we used 225 females who were also tested for the presence of HPV in DNA from cervical smears, and 109 patients with breast cancer. The CIN patients and controls were from a population-based case-control study³.

To analyse the codon-72 polymorphism, we digested a 199-base-pair (bp) product from polymerase chain reaction (PCR) with the restriction enzyme *Bst*UI and separated the fragments by polyacrylamide gel electrophoresis⁴. Only the arginine allele is cleaved, giving two fragments of 113 and 86 bp, whereas the proline allele is not cut (the fragment remains 199 bp long). More than 100 samples have been analysed both by this

method and by constant denaturant gel electrophoresis, giving the same results. Cytological specimens from the CIN patients and from the controls were analysed for HPV using nested PCR and type-specific primers³. The patients with breast cancer were included as additional controls.

Our results on genotype distribution are presented in Table 1. We did not find any significant overrepresentation of homozygotes for the arginine allele either among the cervical-cancer patients or the CIN II–III patients compared with controls. The frequency in patients with breast cancer was similar to the other controls. Comparison of the total patient group with cervical malignancy with the 334 controls, regardless of HPV status, revealed no significantly increased risk for women carrying the Arg/Arg genotype (odds ratio, 1.09; 95% confidence interval, 0.73–1.61; $P=0.74$). The power of this study to detect a twofold increase in the susceptibility to HPV-associated malignancy for Arg/Arg homozygotes was 92%.

HPV-16- or HPV-18-positive cervical-cancer patients revealed an odds ratio of 1.24 (95% confidence interval, 0.51–2.97; $P=0.74$) for the Arg/Arg homozygotes, compared to HPV-positive controls. The probability of detecting a sixfold-increased risk among HPV-positive Arg/Arg homozygous individuals, as found by Storey *et al.*, is 96% in our study. We were therefore not able to confirm that HPV-positive women carrying the Arg/Arg genotype have an increased risk of developing cervical cancer.

The frequencies of the *p53* codon-72 genotypes vary according to ethnic group. The frequency in our control group is similar to that found in a Swedish study⁵. Storey *et al.*¹ report frequencies similar to those found in a Japanese population⁶. As infection with cancer-associated HPV types is relatively common among cytologically normal women as well⁷, other environmental or genetic cofactors are required for cervical carcinogenesis. It may be that the virus load or the status of HPV integration influences the susceptibility to HPV-associated cancers. An association with HLA specificity has been found among both cervical-cancer and CIN patients^{8,9}, and a strong interaction between tobacco smoke and HPV-16 is indicated¹⁰. Further investigation is needed in different ethnic populations to

Table 1 Genotype distribution in *p53* codon-72 polymorphism

	Pro/Pro	Pro/Arg	Arg/Arg
Controls (<i>n</i> =225)	13 (6%)	90 (40%)	122 (54%)
HPV-positive controls (<i>n</i> =29)*	0 (0%)	14 (48%)	15 (52%)
Cervical carcinomas (<i>n</i> =77)†	10 (13%)	23 (30%)	44 (57%)‡
HPV-positive CIN (<i>n</i> =72)§	2 (3%)	29 (40%)	41 (57%)
HPV-negative CIN (<i>n</i> =19)¶	1 (5%)	7 (37%)	11 (58%)
Breast carcinomas (<i>n</i> =109)	6 (5%)	40 (37%)	63 (58%)

*Controls positive for 'high-risk' HPVs (12 controls were not HPV-typed).

†Cervical-cancer patients positive for HPV-16 or HPV-18.

‡Odds ratio was 1.24 for Arg/Arg homozygous cervical-cancer patients compared with controls; $P=0.74$.

§CIN II–III patients positive for 'high-risk' HPVs.

¶CIN II–III patients negative for 'high-risk' HPVs (one CIN patient was not HPV-typed).

1. Hansen, J. & Lebedeff, S. *Geophys. Res. Lett.* **15**, 323–326 (1988).
2. Jones, P. D. *Geophys. Res. Lett.* **21**, 1149–1152 (1994).
3. Nicholls, N. *et al. Climate Change 1995: The Science of Climate Change* (Cambridge Univ. Press, 1996).
4. Gaffen, D. J. & Ross, R. J. *J. Clim.* (in the press).
5. Karl, T. R. *et al. Bull. Am. Meteorol. Soc.* **74**, 1007–1023 (1993).
6. Easterling, D. R. *et al. Science* **227**, 364–367 (1997).
7. Mearns, L. O., Katz, R. W. & Schneider, S. H. *J. Clim. Appl. Meteorol.* **23**, 1601–1613 (1984).
8. DeGaetano, A. T. *J. Clim.* **9**, 1646–1782 (1996).
9. Changnon, S. A., Kunkel, K. E. & Reinke, B. C. *Bull. Am. Meteorol. Soc.* **77**, 1497–1506 (1996).
10. Kalkstein, L. S. & Davis, R. E. *Ann. Assoc. Am. Geogr.* **79**, 44–64

determine the influence of this polymorphism on HPV-associated carcinogenesis.

Åslaug Helland*, Anita Langerød*, Hilde Johnsen*, Anne O. Olsen†, Eva Skovlund‡, Anne-Lise Børresen-Dale*

*Department of Genetics, Institute of Cancer Research, Norwegian Radium Hospital, Oslo, Norway
e-mail: a.l.b.dale@labmed.uio.no

†Section of Epidemiology, Department of Population Health Sciences, National Institute of Public Health, Oslo, Norway

‡Department of Medical Statistics, University of Oslo, Oslo, Norway

1. Storey, A. *et al.* *Nature* **393**, 229–234 (1998).
2. Helland, Å. *et al.* *J. Pathol.* **171**, 105–114 (1993).
3. Olsen, A. *et al.* *Int. J. Cancer* **61**, 312–315 (1995).
4. Ara, S., Lee, P. S. Y., Hansen, M. F. & Saya, H. *Nucleic Acids Res.* **18**, 4961 (1990).
5. Sjölander, A., Birgander, R., Kiveli, A. & Beckman, G. *Hum. Hered.* **45**, 144–149 (1995).
6. Murata, M. *et al.* *Carcinogenesis* **17**, 261–264 (1996).
7. Hildesheim, A. *et al.* *J. Infect. Diseases* **169**, 235–240 (1994).
8. Helland, Å., Børresen, A.-L., Kristensen, G. & Rønningen, K. S. *Cancer Epidemiol. Biomarkers Prevent.* **3**, 479–486 (1994).
9. Helland, Å. *et al.* *Int. J. Cancer* **76**, 19–24 (1998).
10. Olsen, A. O., Dillner, J., Skrdal, A. & Magnus, P. *Epidemiology* **9**, 346–349 (1998).

Storey and co-workers¹ claim that there is an association between a common variant of the *p53* tumour-suppressor gene and the development of invasive cervical carcinoma. Here we present evidence to refute this, based on a reassessment of the importance of the polymorphism at codon 72 in the *p53* gene for the development of cervical cancer.

We genotyped a large set of both *in situ* and invasive squamous-cell cervical cancer cases and controls. Our material was derived from a strictly population-based epidemiological study of women with a diagnosis of cervical cancer *in situ* ($n=488$), as well as age-matched controls ($n=626$). We also analysed samples from 63 cases of invasive cancer from the same population. All *in situ* cases were collected in the county of Uppsala, Sweden, and only women born in Sweden were included².

We found no statistically significant differences in the distribution of *p53* genotypes between the control women and patients with either *in situ* or invasive cervical cancer (Table 1). Homozygosity for argi-

nine at residue 72 was not associated with an increased risk for *in situ* (odds ratio, 1.06; $P=0.648$) or invasive cancer (odds ratio, 1.12; $P=0.684$). We also compared the frequencies of *p53* codon-72 genotypes, using only cases and controls with a positive human papilloma virus (HPV)-16 history, as determined by amplification of DNA from archival smears using the polymerase chain reaction (PCR), and found no significant difference between them (Table 1). Thus, there is no indication that the *p53* genotype increases the risk of developing cervical cancer in HPV-16-exposed women.

There are at least four possible explanations for the discrepancy between our results and those of Storey *et al.*¹ First, their results could be due to chance. Our study was based on a sample size more than ten times larger than theirs and so gives a more accurate estimate of the association between the *p53* polymorphism and cervical carcinoma. This allows us to detect a true odds ratio of two for association with the arginine-homozygote genotype with more than 95% certainty.

Second, selection bias may be introduced when a convenience sample of blood donors is used as a reference group. By contrast, our study was population based, so the controls were representative of the population in which the cases arose.

Third, the *p53* polymorphism might be relevant only for women from certain populations. This would require that variants of the E6 protein differ in their ability to degrade *p53*, and that geographical differences exist in their distribution. But, given the similarity of British and Swedish populations, ethnic differences are unlikely to explain the discrepancy.

Finally, the DNA source and the techniques used for genotyping might have affected the results. Poor-quality DNA, such as that derived from formalin-fixed tissue, can inhibit PCR amplification, with failure being correlated with the length of the fragment³. Storey *et al.*¹ used a PCR product for the proline allele that is 25% longer than the product for the arginine allele. This length difference may cause a bias in the

PCR in favour of the arginine allele and an overestimate of the number of homozygotes for arginine 72. Such a bias would have less influence on large pieces of DNA, such as that isolated from the blood of the controls. It is dangerous to use an allele-specific assay in a study when the quality of the DNA differs between patients and controls.

Agnetha M. Josefsson*, Patrik K. E. Magnusson*, Nathalie Ylitalo†, Pernilla Quarforth-Tubbin*, Jan Pontén*, Hans Olov Adami†, Ulf B. Gyllenstein*

*Department of Genetics and Pathology, Unit of Medical Genetics, Biomedical Centre, University of Uppsala, Box 589, SE-751 23 Uppsala, Sweden
e-mail: ulf.gyllenstein@medgen.uu.se

†Department of Medical Epidemiology, Karolinska Institutet, Box 281, SE-17177 Stockholm, Sweden

1. Storey, A. *et al.* *Nature* **393**, 229–234 (1998).
2. Ylitalo, N. *et al.* *Int. J. Cancer* (in the press).
3. Lindqvist, A.K. *et al.* *Genome Res.* **6**, 1170–1176 (1996).

Storey and co-workers¹ have reported data suggesting that individuals homozygous for arginine at residue 72 of *p53* (*p53Arg*) are about seven times more susceptible to invasive cervical cancer than individuals who carry at least one proline at that position (*p53Pro*)¹. These preliminary data were supported by *in vitro* evidence demonstrating that the E6 oncoprotein of human papilloma virus (HPV) degrades *p53Arg* more efficiently than *p53Pro*. We have now tested specimens from a total of 1,309 women in three studies for *p53* polymorphisms. We find that *p53Arg* is not associated with an increased risk of preinvasive or invasive cervical neoplasia; indeed, there is a tendency for *p53Arg* to be associated with a decreased risk of neoplasia.

Participants were selected from three projects sponsored by the National Cancer Institute: (1) a 10,000-woman population-based cohort in Guanacaste, Costa Rica²; (2) a 24,000-woman cohort in Portland, Oregon, USA³; and (3) a 529-woman multicentre study of histological subtypes of cervical neoplasia in the eastern United States. We used a test based on the polymerase chain reaction (PCR) for *p53* codon-72 polymorphisms (PCR primers were: *p53.5*, 5'-gaagaccaggtccagatga-3'; and *p53.3*, 5'-ggtaggtttctggtgaagg-3') and single-strand-conformation polymorphism (SSCP) analysis to differentiate between alleles encoding *p53Arg* and *p53Pro* (ref. 4). Direct sequencing of 60 randomly selected specimens confirmed our SSCP findings.

In Costa Rica, women diagnosed with preinvasive cervical lesions (known as LSIL and HSIL, for low- and high-grade squamous intraepithelial lesions, respectively) or invasive cervical cancer had an approximately 2–4-fold smaller risk of disease if they were homozygous for *p53Arg* com-

Table 1 Distribution of codon-72 *p53* genotypes in patients and controls

	Controls		Cancer <i>in situ</i>	Invasive cancer	
	<i>n</i> (%)	Expected (%) [*]	<i>n</i> (%)	OR (95% CI) [†]	OR (95% CI) [†]
All samples					
Pro/Pro	69 (11)	(9)	41 (8)	0.77 (0.50–1.18)	5 (8) 0.74 (0.27–2.02)
Pro/Arg	246 (39)	(42)	191 (39)	1.0 (reference)	24 (38) 1.0 (reference)
Arg/Arg	311 (50)	(49)	256 (52)	1.06 (0.83–1.36)	34 (54) 1.12 (0.65–1.94)
HPV-16-exposed only					
Pro/Pro	12 (11)	(8)	27 (9)	0.77 (0.36–1.66)	n.d. n.d.
Pro/Arg	41 (36)	(41)	120 (39)	1.0 (reference)	n.d. n.d.
Arg/Arg	61 (54)	(51)	159 (52)	0.89 (0.56–1.41)	n.d. n.d.

DNA from controls and *in situ* cases was extracted from cervical smears; that from cases with invasive cervical cancer was from formalin-fixed biopsies. Genotypes were determined by PCR amplification of a 164-base-pair fragment, followed by cleavage of the PCR product by the restriction enzyme *Acl*. n.d., not determined. Detailed methods are available from the authors.

^{*}Expected binomial (Hardy-Weinberg) proportions.

[†]Odds ratios (OR) with 95% confidence intervals (CI) were calculated.

Table 1 Distribution of and risk associated with p53 codon-72 polymorphisms

Study group	n	p53 Pro/Pro (%)	p53 Arg/Pro (%)	RR-Arg/Pro (95% CI)*	p53 Arg/Arg (%)	RR-Arg/Arg (95% CI)*
Costa Rica study						
Population†	123	6.5	48.8		44.7	
LSIL	81	16.1	55.6	0.46 (0.18, 1.2)	28.4	0.26 (0.09, 0.7)
HSIL	117	12.8	45.3	0.47 (0.19, 1.2)	41.9	0.48 (0.09, 1.2)
Cancer	49	14.3	44.9	0.42 (0.14, 1.3)	40.8	0.42 (0.13, 1.3)
Portland study						
HPV(−) normal	109	3.7	29.4		67.0	
HPV(+) normal	108	8.3	38.0	0.57 (0.16, 2.0)	53.7	0.35 (0.10, 1.2)
LSIL	93	5.4	31.2	0.73 (0.18, 3.0)	63.4	0.65 (0.17, 2.5)
HSIL	115	9.6	35.7	0.47 (0.14, 1.6)	54.8	0.31 (0.10, 1.0)
Eastern United States study (squamous-cell carcinoma)						
Control	245	8.2	40.0		51.8	
HSIL	47	6.4	40.4	1.30 (0.35, 4.8)	53.2	1.30 (0.36, 4.8)
Cancer	88	10.2	40.9	0.82 (0.34, 2.0)	48.9	0.75 (0.32, 1.8)
Eastern United States study (adenocarcinomas and rare histological types)‡						
Control	245	8.2	40.0		51.8	
AIS§	35	5.7	28.6	1.0 (0.21, 5.0)	65.7	1.80 (0.40, 8.3)
Cancer	99	9.1	48.5	1.1 (0.46, 2.6)	42.4	0.74 (0.31, 1.7)

*p53 Pro/Pro was considered as the baseline when calculating estimates of relative risk (RR); 95% confidence intervals (CI) were calculated.

†Random sample of the Costa Rican (Guanacaste) population.

‡Includes 93 pure adenocarcinomas, 22 adenosquamous carcinomas and 19 other rare histological types.

§AIS, adenocarcinoma *in situ*.

pared with women homozygous for p53Pro (Table 1). The risk for heterozygotes was also roughly halved. These findings were reproduced in our Portland study, in which women who were heterozygous or homozygous for p53Arg had a 1.5–3-fold decreased risk of disease (Table 1).

Cervical cancers with a glandular component are known to be more strongly associated with infection by HPV-18 than are the more common squamous-cell carcinomas of the cervix⁵, and Storey *et al.*¹ focused on the HPV-18 E6 protein. We therefore assessed the effect of p53 polymorphisms on disease risk by looking at tumour histology in our third case-control study of squamous cell tumours and adenocarcinomas of the cervix. We found no evidence that women with p53Arg were at increased risk of disease (Table 1).

There was no indication either of a positive association between p53Arg and cervical neoplasia in analyses restricted to participants who were positive for HPV-16. Analysis of pooled data from the three studies restricted to HPV-18 suggested only a weak, non-significant association between p53 status and cervical neoplasia, despite the 79% power of our study to detect a risk factor of seven, the estimate of Storey *et al.*¹. Among the 67 women positive for HPV-18, the prevalence of p53Arg was 40% among cytologically normal individuals ($n=10$), 59% among LSIL patients ($n=17$), 60% among HSIL patients ($n=20$), and 60% among cancer patients ($n=20$; $P=0.72$). The comparable proportion of HPV-18-negative controls carrying p53Arg was 54%.

Longitudinal data from our Costa Rican cohort were also analysed ($n=75$ women; median follow-up time was 44 months). We tested whether carriers of p53Arg had an

altered risk of incident HSIL or of persistent LSIL during follow-up, but found no evidence for a detrimental effect caused by p53Arg (data not shown).

Given the consistent results from these three studies, it is unlikely that p53Arg is associated with any increase in susceptibility to cervical neoplasia.

Allan Hildesheim*, Mark Schiffman*, Louise A. Brinton*, Joseph F. Fraumeni Jr*, Rolando Herrero†, M. Concepcion Bratti‡, Peter Schwartz§, Rodrigue Mortel||, William Barnes¶, Mitchell Greenberg#, Larry McGowan☆, David R. Scott**, Maureen Martin††, Jesus E. Herrera††, Mary Carrington††

*Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6130 Executive Boulevard, Bethesda, Maryland 20892, USA

†International Agency for Research on Cancer, 69372 Lyon, France

‡Caja Costarricense de Seguro Social, San Jose, Costa Rica

§Yale University School of Medicine, New Haven, Connecticut 06519, USA

||Milton S. Hershey Medical Center, Hershey, Pennsylvania 17033, USA

¶Georgetown University Lombardi Cancer Center, Washington DC 20007, USA

#Graduate Hospital, Philadelphia, Pennsylvania 19146, USA

☆George Washington University Medical Center, Washington DC 20037, USA

**Kaiser Permanente, Portland, Oregon 97227, USA

††Intramural Research Support Program, SAIC-Frederick, National Cancer Institute-FCRDC, Frederick, Maryland 21702, USA

1. Storey, A. *et al.* *Nature* **393**, 229–234 (1998).

2. Herrero, R. *et al.* *Pan-Am. J. Public Health* **1**, 362–374 (1997).

3. Hildesheim, A. *et al.* *Cancer Epidemiol. Biomarkers Prev.* **7**, 1035–1041 (1998).

4. Cullen, M. *et al.* *Am. J. Hum. Genet.* **60**, 397–407 (1997).

5. Kjaer, S. K. & Brinton, L. A. *Epidemiol. Rev.* **15**, 486–498 (1993).

Storey *et al.* reply — These reports assess the frequency of the p53Arg allele in different populations and conclude that homozygous p53Arg is not a risk factor for cancer associated with human papilloma virus (HPV). The functional differences between the p53 isoforms that we have described^{1,2} provoked our initial epidemiological study. As we concluded then, it is crucial that investigations should be extended to different populations, and we are encouraged that such studies are underway.

There are several potential reasons for the apparent discrepancies between our original study and those reported here. Our study comprised a smaller number of samples from a different population. In addition, rather than leukocyte DNA, we screened microdissected tumour material, in which we demonstrated that it was always the proline allele that was lost in cases of loss of heterozygosity. Furthermore, it is important to compare the different screening methodologies that have been used as this may also be a source of incorrect allelic assignment, which would lead to a bias towards a null hypothesis.

In such studies, the composition of the control population is a major factor affecting the outcome. We are now concentrating our efforts on populations with a higher frequency of the proline allele. Preliminary results from a Brazilian population strongly support our original observations. In addition, when the control population consists of cytologically normal but HPV-positive individuals, the association of the p53Arg allele with the risk of tumour development is greatly increased; a similar trend is reported here by Hildesheim *et al.* All of these factors need to be taken into consideration in any future studies of this polymorphism.

Continuing exploration of the differential effects of this polymorphism should further unravel the complexities of p53 function.

Alan Storey*, Miranda Thomas†, Ann Kalita‡, Catherine Harwood*, Daniela Gardiol†, Fiamma Mantovani†, Judith Breuer§, Irene M. Leigh†, Greg Matlashewski‡, Lawrence Banks†

*Imperial Cancer Research Fund, Skin Tumour Laboratory, 2 Newark Street, London E1 2AT, UK

†International Centre for Genetic Engineering and Biotechnology, Padriciano 99, I-34012 Trieste, Italy

‡Institute of Parasitology and McGill Cancer Centre, McGill University, Macdonald Campus, 21,111 Lakeshore Road, Ste Anne de Bellevue, Québec H9X 3V9, Canada

§Department of Medical Microbiology, St Bartholomew's and The Royal London Hospital School of Medicine and Dentistry, Queen Mary and Westfield College, London E1 1BB, UK

1. Storey, A. *et al.* *Nature* **393**, 229–234 (1998).

2. Thomas, M. *et al.* *Mol. Cell. Biol.* (in the press).

The riches found in a prospector's pan

Dorothy Hodgkin: A Life

by Georgina Ferry
Granta: 1998, 419 pp. £20

Judith A.K. Howard

Dorothy Hodgkin: A Life is a fascinating and inspiring biography of the only British woman to win the Nobel prize. Its contents are as accessible to the non-scientist as to those of us who knew Dorothy Hodgkin and her scientific work well. Having known Dorothy made the whole book a delightful re-run of much familiar history, but it also introduced some new stories or fascinating new slants on old stories. Not having known her, you would surely feel that you had done so after travelling through these thoroughly researched pages, which skilfully condense an enormous amount of detail and data.

The book starts with Dorothy's background: her grandparents and parents and early years at school. Her father, John Crowfoot, was a civil servant working in the Egyptian Education Service in Cairo when Dorothy was born in May 1910. With the outbreak of the war in 1914, her mother, Molly, decided to take the girls (Dorothy by now had two sisters, Joan and Betty; a third, Diana, was born towards the end of the war) back to England with their nurse, Katie. Molly soon rejoined her husband in Egypt, and, for the rest of the war, the Crowfoot girls were looked after by Katie and their grandparents.

Dorothy has attributed much of her independent spirit to this period in her childhood, when learning self-reliance seemed as natural as it was essential, and it was not regarded as neglectful for her parents to leave the girls 'alone', since many families were thus separated by the difficulties of war. This chapter has some lovely recollections of her early childhood abroad and the thrill she felt in basic discoveries, the expeditions and excavations and many travels with her parents.

It was during one long visit to Khartoum that Dorothy was introduced to a family friend known as 'Uncle Joseph', who was working in the new Wellcome Laboratory there. To amuse the girls one day, he showed them how to 'pan' for gold, and on trying this later in their garden stream, Dorothy discovered some shiny black material, which she analysed with Uncle Joseph's help and showed to be "ilmenite, a mixed oxide of iron and titanium; and titanium was not in my school chemistry book".

Impressed by Dorothy's enthusiasm, Uncle Joseph gave her an old portable surveyor's box when she left to return to England; it contained the basic chemicals needed for simple analyses, which she was able to continue with great joy in her 'laboratory' at



Something to celebrate: Dorothy with her husband Thomas at the Nobel ball.

home in the attic at Geldeston. This is described wonderfully in the opening paragraph of the book, where the whole scene is set captivatingly by Ferry. Dorothy had already started to learn some chemistry before this trip and was fortunate to have a gifted and encouraging young female chemistry teacher, Criss Deeley, at the Leman School in Beccles.

This chapter lays the foundation for the career that follows. Ferry describes how it was then to be a young woman at Oxford's Somerville College in the late 1920s, when it was exceedingly rare for women to study science at all, and to be invited to carry on with research work thereafter, a daunting prospect for Dorothy.

Largely on Uncle Joseph's advice, Dorothy went to Cambridge in 1932 to work with the young and politically controversial new appointee in X-ray crystallography, J. D. Bernal (known to everyone as 'Sage'). The excitement that flowed between them and in the laboratory is obvious from the way Dorothy writes about her work at this time. It was, she says, "rich with new discoveries", as they saw for the first time the impact this new subject was to make on biological science, and they began the task of obtaining, or growing, crystals of biologically important compounds and taking their X-ray photographs, although the complete structure solutions they knew to be frustratingly beyond the techniques available to them at that time.

Ferry manages to convey, in her imaginatively descriptive prose, some of the huge delight and fascination that Dorothy herself must have felt in these early years of studying crystals and their structures, and indeed,

Dorothy never really lost this enjoyment throughout her career.

It might seem to the non-scientist that there is much crystallographic detail in parts of the book, but Ferry has cleverly diluted many of the more heavily scientific sections with lighter reports and quotations from Dorothy's own biographical material, and the theme would not be lost if certain parts were to be skipped by any less assiduous readers, who were only marginally interested in the niceties of X-ray diffraction.

Dorothy had barely started her research for her doctorate in Cambridge when Somerville offered her a fellowship back in Oxford. She was torn between the very tempting offer of a job and her passion to continue the work she had begun with Sage. Somerville shrewdly suggested she continue to work in Cambridge for the first year and then devote most of the second year to research in Oxford. This offer "amazed and delighted" Dorothy and so she accepted. Sad though she was to leave Sage, they continued to exchange ideas and to work together, and they kept in touch scientifically for the rest of his life.

It was shortly after her return to Oxford that she met Thomas Hodgkin, whom she married in December 1937. The story of meeting Thomas and their love for one another is sensitively illustrated with many letters between them, and it seems incredible that Dorothy was required to resign her fellowship at Somerville on marriage, although she was reinstated automatically afterwards. Her work on the structure of insulin had already started by now and this story is interwoven with her early days in Oxford and the rapidly growing interest in protein structure.

Dorothy gave birth to their first son,

Luke, a year after they were married and shortly afterwards suffered a breast abscess that simply wouldn't heal. This resulted finally in an operation to lance the abscess, but by the following Easter (1938), Dorothy was in pain all over and severely restricted in her movements. She had suffered an acute attack of rheumatoid arthritis at the age of 28 and the consequences of this never left her. She was sent to the Spa at Buxton, Derbyshire, for treatment that gave her some remission, but such 'cure' as she knew was never to be permanent. For many years, Dorothy showed great courage in the face of personal physical suffering. This started for her at an age when most young women were settling into family life and not struggling, in addition, with the enormously challenging problems of an academic career, newly opened in X-ray crystallography.

The following chapter describes her very important work on penicillin during the war, the birth of their second child, Elisabeth, their second son, Toby, and her election to the Royal Society in 1947. Kathleen Lonsdale, a fellow crystallographer from University College London, was one of the first two women elected to the society just two years previously. Dorothy's election brought many congratulations from the crystallographic world and the pithy comment from Justin Waddington, "If a woman FRS can have three children, anyone might do anything".

The next chapters relate the details and personal stories surrounding the study and solutions of the structures of vitamin B₁₂ and insulin. It was during this period that the electronic computer made a vital appearance in X-ray crystallography. It was to revolutionize the field and in particular to speed up the very tedious manual calculations previously made by Dorothy using the famous Beevers-Lipson strips, a brilliant innovation from Arnold Beevers and Henry Lipson, two UK crystallographers based in Bragg's laboratory in Manchester. It was for her out-

standing work on vitamin B₁₂ and "other important molecules of biological significance" that Dorothy was awarded the Nobel prize for chemistry in 1964. She received the wonderful news in Ghana, where she was staying with Thomas at the time. He had been appointed in 1961 as the first director of the new Institute of African Studies at the University of Ghana and Dorothy tried to spend some time with him there each year. But following the military coup in 1966, Thomas returned to Oxford and a fellowship at Balliol. Soon afterwards, they moved into Crab Mill, Illmington, to the old house of Thomas's parents, where Dorothy was to spend the rest of her life. Like her parents before her, Dorothy had managed her family and marriage 'split' initially between towns in the United Kingdom and then for several years in countries abroad.

The year after receiving the Nobel prize, she was awarded the Order of Merit, the highest British honour any citizen can receive and of which there are only 24 members at any time; new members really do fill dead men's shoes. Dorothy was the first woman to receive this honour since Florence Nightingale in 1907.

Honours followed upon honours as Dorothy went on to receive many honorary degrees and scientific accolades from around the world. She became the president of the Pugwash Conferences on Science and World Affairs in the mid-1970s and campaigned tirelessly on behalf of international peace for many years. On the question of disarmament, she wrote, "I do think that Pugwash has been useful as one more way of making close contacts between leading scientists in different fields, probably in preventing worse from happening in the way of wars, probably assisting, if very imperfectly, the practice of peace". She visited Russia, China and Vietnam when it was very much more difficult for scientists to travel there, and made lifelong friends among the crystal-

lographers in these countries.

Dorothy nominally retired from Oxford University 17 years before she died, but she continued working on the structure of insulin for many more years; she was the chancellor of Bristol University during most of this time, and continued to travel around the world for Pugwash and for other international seminars and meetings, seemingly never tiring of working with and for others, despite her increasingly frail physique, and in particular after Thomas died in 1982.

Ferry has demonstrated her very considerable talents as a science writer in this brilliant biography of Britain's most famous, yet most modest and gentle, woman of science and protagonist of peace. It is a must for all serious crystallographers as well as science historians, since it describes wonderfully well so much of our crystallographic history, both the techniques and the personalities, over a lifetime of very significant discoveries and technological developments. □

Judith A. K. Howard is in the Department of Chemistry, University of Durham, Durham DH1 3LE, UK.

Casting out the printers' devils

The Nature of the Book: Print and Knowledge in the Making

by Adrian Johns

University of Chicago Press: 1998. 753 pp.

\$40, £31.95

Michael Hunter

Francis Bacon saw printing as one of the inventions that had "changed the appearance and state of the whole world", and ever since his time the printed book has been a potent symbol of the preservation and transmission of reliable knowledge. Indeed, the chief landmarks of the history of science are books like Copernicus's *De revolutionibus* (1543) or Newton's *Principia* (1687). Yet only a moment's reflection is required to realize that the publication of science books, as much as any other kind of book, is a complex process, involving a range of parties and interests, not all with common aims.

It is basically this insight, applied to the burgeoning scientific culture of seventeenth-century England, that underlies Adrian Johns's provocative and stimulating, if ultimately not wholly satisfactory, book. He argues that the nature of the book in seventeenth-century England was inextricably bound up with the ways in which books were printed and marketed, and that the ultimate control over what ideas were presented, and the manner in which they reached the public, had far less to do with the authors than with the producers — the printers, booksellers and others involved in the book-trade

Bubbling bauble

Io, the innermost of the four moons orbiting Jupiter, sports active volcanoes on its surface. It orbits so close to its parent planet that the tidal pull of Jupiter squeezes the moon rhythmically, generating heat in its interior which drives the volcanic activity. The image shown here was obtained by the Galileo space probe, and is one of many examples of space imagery to be found in *Empire of the Sun: Planets and Moons of the Solar System* by John Gribbin and Simon Goodwin (Constable, £15.99).



at the time. The result was that the dissemination of reliable data was far more problematic for scientists of the day than has generally been appreciated.

A striking example of this is provided by the first scientific periodical, *Philosophical Transactions*. Within two years of its inauguration in 1665, the run of issues was 'contaminated' by an unauthorized publication in matching format put out by the same printer; this took in at least one early reader, the diarist John Evelyn, who dutifully had the offending item bound into his set of the journal.

Equally strange is the case of John Flamsteed's *Historia Coelestis*, of which two distinct versions were published, one by Flamsteed himself, the other by Isaac Newton and Edmond Halley using Flamsteed's data. The latter differs from the former not least in lacking the star charts that Flamsteed considered crucial to the work. Although this episode owed more to a clash of personalities than to the normal circumstances of publishing at the time, it is nevertheless grist to Johns's mill. He revels in the instability of the printed word that such cases reveal, arguing (perhaps with some exaggeration) that Newton, Boyle and their contemporaries spent as much time negotiating how their books were printed and read as in experimenting.

By way of background, a series of detailed chapters takes us on a fascinating exploration of the world of books in early modern London, the points made being reinforced at every stage by the profuse use of well-selected illustrations. We encounter the printers, the booksellers' shops and the Stationers Company through which the trade in printed material was organized. We also learn about the procedures for censorship that existed, and the way in which they operated. Here, Johns repeatedly emphasizes how fully the producers of books were in control; indeed, the author appears in a very peripheral role, dependent on the whims of printers when consulted at all.

Yet, although Johns makes clear what a cut-throat business it was, he also illustrates the producers' awareness of the need for some form of supervision over what was printed and by whom. From the sixteenth century onwards this was achieved by a mutually advantageous arrangement between the Stationers and the prerogative powers of the Crown. But potential for conflict existed, not least in the aftermath of the Civil War, as comes out particularly through two chapters dealing with the attempt by John Streater and Richard Atkins to use royal power to subvert the Stationers' privileges, with rival histories of printing used to support the different sides in this dispute.

This leads to perhaps the most fascinating chapter in the book, that on "the Physiology of Reading", in which Johns brings together the history of science and cultural history. Deploying the views on perception of early



The written word conveyed in oils: painting by the Dutch seventeenth-century artist Edwaert Colyer.

modern authors like Thomas Willis, who emphasized the interconnection of vision and imagination, Johns illustrates how these resonated with contemporary theories of the 'passions', understood as impulses generated either by sensations or by internal bodily states, to produce an awareness that the encounter between the reader and the printed word was radically unstable. Only careful supervision and control could overcome this, and here, as elsewhere, Johns wants to argue for the role of trust.

This was crucial precisely because the book itself was not a self-evidently reliable commodity; readers therefore had to know how to approach any text through knowledge of its background in order to establish its trustworthiness. Moreover, he wants to argue that in an important sense this relativism of the printed word transcends the specific context of early modern practice. Responding to the argument that the invention of the steam press in the nineteenth century brought about a real revolution that had simply failed to occur earlier, he argues that, on the contrary, many of the issues of trust and credibility that he illustrates in his book are still with us, in the context not only of print but also of electronic media.

In making his case, Johns has evidently come to feel that the most effective method is through a detailed exposition of the circumstances of the production and reception of books in seventeenth-century England, and particularly through recounting key episodes at length. Certainly, this is something for which he has a real flair, with an enviable eye for telling detail, a skill in narrative and an ability to intersperse this with ingenious, often deflationary, asides. The question is whether he has allowed himself to be slightly carried away. Overall, the book is 750 pages long, while the chapters are massive — the longest runs to

129 pages — and the thread of the argument sometimes gets lost in them. Moreover, the cumulative effect is wearying: only on page 444 do we reach the chapters devoted to the Royal Society and science books.

This is the more problematic since, although detailed and informative, the book is suggestive rather than definitive. There are many aspects of the history of the book in seventeenth-century England that are not pursued here. Johns's concern is essentially about the control of the printed word, and he has virtually nothing to say about print runs and the relative popularity of different kinds of books at the time. (It is worth noting that his critique of claims for a 'print revolution' in the period focuses almost exclusively on the fixity of print rather than its ability to produce massive numbers of copies of a book in the time it took scribal copyists to produce one.)

But there are also omissions directly germane to Johns's chosen theme. The book focuses almost exclusively on London, and publishing in Oxford is mentioned only in passing: this means that there is no account of the attempts to make learned publishing viable at the Clarendon Press under the auspices of Bishop John Fell in the late seventeenth century, although these have interesting resonances with the efforts of the Royal Society in London which are dealt with at length.

Even on topics that *are* covered, Johns does not always have the final word. For instance, although there is an acute and interesting section on Robert Boyle, dwelling particularly on his complaints about unauthorized editions of his works at home and abroad, no attempt is made to test Boyle's statements by surveying the actual publishing history of his works. Yet this would have revealed a rather different picture, as will be documented in detail in the new edition of his works due to be published next year.

For one thing, Johns is so keen to make Boyle's experiences with 'usurpation' seem normal in the cut-throat publishing world of the time that he accepts his complaints at face value, failing to consider the possibility that they might form part of a pattern of 'protesting too much' on Boyle's part — in relation not just to booksellers and literary thieves, but also to every other aspect of his life as an author — which arguably tells us less about the actual state of affairs than about a slightly obsessive streak in Boyle with significant implications for an understanding of his intellectual personality. More importantly, Johns seems unaware that Boyle in fact adopted a rather canny publishing practice, always working with more than one bookseller at a time in an evident attempt to play them off against one another — a strategy that makes all the better sense against the rather unstable commercial background sketched here.

Such instances could be multiplied, and the fact is that, for all its length, *The Nature of the Book* lacks not only definitiveness, but also balance: throughout its 750 pages it is arguing a case rather than saying the final word even on scientific publishing in seventeenth-century England.

Yet Johns's decision to write a big book rather than a smaller one may be germane to the very issues that dominate the work itself, namely the extent to which people's approach to the printed word is affected by factors extrinsic to a book's actual content. People defer to long books, and Johns may well be correct in thinking that a weighty tome like this will gain an authority for his iconoclastic view of print that would not have been the case with a work half the length. Yet, if he is right in this, it is to be hoped that he will not be a victim of his own success: readers should not be discouraged by the sheer length of the book from sampling some of the lively and insightful sections that make it up. □

Michael Hunter is in the Department of History, Birkbeck College, University of London, London WC1 7HX, UK.

Roaming the landscape

Fractal River Basins: Chance and Self-Organization

by Ignacio Rodríguez-Iturbe and Andrea Rinaldo
Cambridge University Press: 1997. 547 pp.
£70, \$100

David R. Montgomery

The search for meaning in the branching patterns of channel networks has seduced geomorphologists for decades, and *Fractal River Basins* makes a compelling case for a fundamental order in landscape form. Ignacio Rodríguez-Iturbe and Andrea Rinaldo provide a benchmark synthesis and passionate exploration of the ways in which fractal geometry can be used to analyse river networks, and how the concepts of optimal channel networks (OCNs) and self-organized criticality (SOC) can be applied to models of landscape evolution. In a rigorous and engaging style, they explore the themes of necessity and chance in river-basin evolution to develop the argument that fractal geometry contributes more to geomorphology than just a convenient shorthand for the thumbprint of river incision.

Fractal River Basins goes beyond the descriptive application of fractals to examine the universal aspects of landscape development. An excellent overview of river-basin morphology opens the book and leads into a discussion of the use of fractals to describe the wide variety of scale-free aspects of river

networks and river-basin morphology.

The authors delve into examples where fractal descriptions and models are supported by observations from real networks, but they also discuss how channel networks are not simple, self-affined monofractals, and emphasize that planform models of network structure are not demanding when it comes to statistical tests of their veracity. A wide variety of models produce arborescent patterns that resemble natural channel networks, and the robust generality of the fractal aspects of channel networks has led many geomorphologists to question the utility of fractals for elucidating the processes of network growth and landscape evolution.

Rodríguez-Iturbe and Rinaldo argue that models of landscape evolution should be assessed by examining the three-dimensional structure of basins (that is, including elevation), and they propose that OCNs provide a bridge between fractal patterns and landscape-forming processes. In a discussion that summarizes a decade of cutting-edge research, they start from the premiss that channel networks are characterized by minimum rates of energy expenditure to show that the connectivity and elevation structure of OCNs resembles that of natural drainage basins.

They argue that both chance and necessity govern river-basin evolution — chance through the effect of random variations in initial conditions and necessity through minimization of the total rate of energy expenditure both for the whole network and locally throughout a river network. In drawing previous work together into a single chapter the authors provide a very welcome and valuable synthesis of a wide range of studies on OCNs and a comprehensive case for their relevance to river-basin evolution.

The authors' examination of how the structure of river basins is formed by SOC contrasts with a perception among some process geomorphologists that SOC offers little to those seeking to understand land-form. It is not too surprising, therefore, that in framing their discussion of landscape self-organization and river-basin evolution, Rodríguez-Iturbe

and Rinaldo decry the "reductionist, process-oriented tenet to which most geomorphology is committed". But both reductionism and the synthetic examination of emergent properties are necessary tools for the well-rounded geomorphologist, given that process geomorphology in general seeks to understand how and to what degree the processes driving landscape evolution translate across a wide range of scales.

Moreover, some process-based understanding is necessary to evaluate the implications of the principle of minimum energy dissipation; without a tie to landscape-forming processes, the existence of broad, ubiquitous, self-similar characteristics of landscapes teaches us little about how particular landscapes evolve. Even though a fundamental tension underlies the search for universal and landscape-specific aspects of land-form, there is value in searching for insight in both the ubiquitous and the discriminating characteristics of landscapes.

However one looks at the subject, *Fractal River Basins* is an important book that explores the fundamental dynamics behind the generally uniform appearance of river networks and fluvially sculpted landscapes. The reader seeking an exploration of the relevance of fractal geometry to the description of river networks will be rewarded with a state-of-the-art review.

The seductive beauty of channel networks will continue to inspire investigators to examine the deep symmetry expressed in landscape form, and *Fractal River Basins* provides a bold synthesis of probability, physics and geometry that will stand as a definitive work on a complex subject. Geomorphologists and physicists alike will find the book thought-provoking, and I highly recommend this stimulating work as a means of focusing one's thinking on the role of fractals, OCNs and SOC in fluvial geomorphology and landscape evolution. □

David R. Montgomery is in the Department of Geological Sciences, University of Washington, Seattle, Washington 98195-1310, USA.

The splendour of Yellowstone

This aerial view of Sour Creek, showing the mosaic of meandering streams, ponds and wetlands across open pasture, is one of many photographs taken by Norbert Rosing for *Yellowstone* (Firefly, \$24.95), in which the photographer aims to convey the fascination of Yellowstone National Park, with its enduring geological monuments, geysers, plants and wildlife.



Quantum non-demolition measurements in optics

Philippe Grangier, Juan Ariel Levenson & Jean-Philippe Poizat

Quantum non-demolition measurements are designed to circumvent the limitations imposed by Heisenberg's uncertainty principle when performing repeated measurements of quantum states. Recent progress in quantum optics has enabled the experimental realization of quantum non-demolition measurements of the photon flux of a light beam. This achievement bears on fundamental issues about the ultimate sensitivity of measurements, and may open the way for applications such as noise-free information tapping in optical telecommunications.

It is written at the very heart of quantum mechanics that a precise measurement in the microscopic world is not possible without the introduction of a perturbation or 'back action' inherent to the very fact of measurement. This principle, which has been known since the 1930s¹, can be directly related to Heisenberg's well known uncertainty relations. Using the fact that the quantum formalism describes physical quantities as non-commuting operators (that is, as mathematical objects A , B such that $AB \neq BA$), the Heisenberg inequalities state that the product of the dispersions of (the 'uncertainty' in) A and B has a lower bound: $\Delta A \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|$. Therefore, for non-commuting operators, a very precise measurement of A , resulting in a very small dispersion ΔA , will be associated with a large value of ΔB . Although this does not restrict directly the precision in the measurement of A itself, the large fluctuations induced in B may eventually couple back to A , which will then also be perturbed. This 'measurement back action' has far-reaching consequences from a practical point of view, because it may prevent the retrieval of the initial result in a series of repeated measurements. In response to this problem, Braginsky, Thorne, Unruh, Caves and others introduced in the 1970s the concept of "quantum non-demolition" (QND) measurement²⁻⁸, in which a measurement strategy is chosen that evades the undesirable effect of back action. The key issue is to devise measurement schemes in which the back-action noise is kept entirely within unwanted observables, without being coupled back onto the quantity of interest. This quantity then remains uncontaminated by the measurement process, allowing repeated measurements to be performed with arbitrary high accuracy.

Originally, QND ideas dealt with mechanical oscillators designed for detecting gravitational waves^{9,10}, so-called Weber bars^{11,12}. But these devices appeared to be limited by technical difficulties, and QND moved on to other fields. In the mid-1980s, the emerging field of quantum optics¹³⁻¹⁵ seemed to be particularly well suited for implementing QND measurements¹⁶⁻¹⁸. The main reason for this was that the technical quality of optical sources and detectors was good enough for sensitivity at the quantum noise level to be achieved. Moreover, we show below that to manipulate the quantum noise of a light beam requires techniques of nonlinear optics that have also developed rapidly in recent years. The first achievements of experimental quantum optics were the generation of individual beams with reduced quantum noise^{19,20}. It took longer for the first QND schemes to be successfully implemented²¹⁻²⁷, because they require two beams and a more sophisticated characterization procedure, as we shall explain below.

To summarize some of the basic features of quantum measurements, we use the standard textbook example of a quantum measurement: the Stern-Gerlach experiment. We then show how the effectiveness of a QND measurement can be evaluated in the context of quantum optics. Finally, we present the latest experimental results, which exploit either third- or second-order optical

nonlinearities, giving emphasis to the realization of repeated QND measurements. The ideas developed from QND can be applied for processing information-carrying light beams, using 'noiseless' amplifying systems.

Basic principles

Consider the measurement of a particle's spin using the Stern-Gerlach (SG) apparatus (Fig. 1). It is usually stated in textbooks that an SG experiment 'measures' the spin of the particle. But let us ask the question: what properties are actually required in order to qualify this as a good measurement of spin?

The first requirement is obvious: the device should be able to distinguish between different values of the spin variable. As the measured quantity can assume only discrete (quantized) values (here $\pm \hbar/2$), it is enough that these eigenvalues be well resolved. In a more general case, with many closely spaced eigenvalues or a continuous spectrum, one can easily specify a required degree of accuracy.

As a second requirement, it is reasonable to expect that repeated measurements give the same result. For instance, in the SG filter shown in Fig. 1b, particles with a chosen spin value $+\hbar/2$ go through, and the others ($-\hbar/2$) are blocked. Then the transmitted particle will again go through any other SG filter with the same orientation. As a result of this property, it is usually said that an SG filter is a good 'quantum-state preparation' (QSP) device.

For the third requirement, which is crucial for our purpose, let us look at the following situation: we wish to monitor an external perturbation acting on the particle between two successive measurements, through the changes that it induces on the spin state on the particle (Fig. 1b). In that case, it is clear that the QSP property is not enough: any spin flip of the particle from $+\hbar/2$ to $-\hbar/2$ will cause its 'demolition', because it will be filtered out by the next SG device, interrupting the monitoring. At this point the most specific aspect of a QND measurement enters: the non-demolition property itself. Whereas a filter is designed to select a quantum state with a specific prescribed eigenvalue, a QND device should be able both to measure and to keep the particle going, whatever the result of the measurement. Therefore, the SG filter quoted above can be efficient for measuring and preparing a quantum state, but it is not a QND device, precisely because preparing the state $+\hbar/2$ is done by demolishing the state $-\hbar/2$. We will see below how to improve on this.

In this example, it is noticeable that without any external perturbation the spin does not evolve between two successive measurements. This makes the monitoring easier, and it is said that the spin is a 'good' QND variable. More serious problems appear when the free evolution couples the measured quantity to another one that carries very large measurement-induced fluctuations due to Heisenberg's inequality, as explained above. Then the very possibility of the monitoring is made impossible by the back

action of this measurement noise. This difficulty has been discussed extensively⁶, especially in the context of position measurement for a particle. The general conclusion is that the monitoring is not always possible: it has to be done on carefully chosen QND variables, the most obvious ones being simply constants of motion (like the spin in our example).

Once such a quantity is identified, a difficult experimental issue shows up: a measurement coupling must be designed that is able to perform an accurate measurement (that is, leave the system in a known quantum state) without ever demolishing the state. As an example, let us look again at the SG case: how can the spin be measured in a non-destructive way? Once the trajectory has been split by the magnet, one should identify the path followed by the particle, again in a non-destructive way. But this just shifts the difficulty one step further along, and a much better solution is actually to abandon the SG apparatus and to use instead a so-called 'indirect' measurement: employing another particle, which undergoes a spin-dependent interaction, to 'read out' the information about the first particle's spin state (Fig. 1c).

This brings us to the standard way in which QND measurements

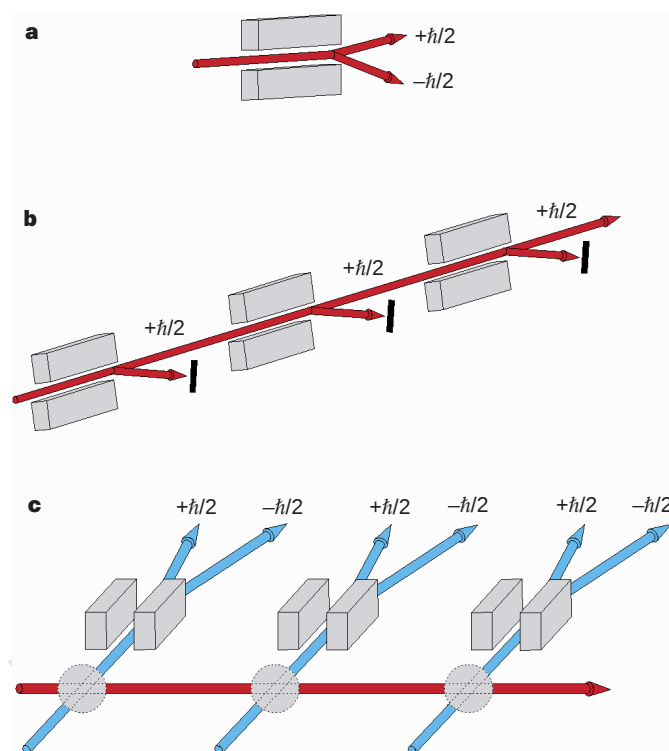


Figure 1 Simple examples of quantum measurements. **a**, Spin measurement using a Stern–Gerlach (SG) apparatus. The particle is deflected in a direction that depends on the value of its spin component along the direction of the magnetic field. Then the spin is ‘measured’ by detecting the position of the particle after some free propagation. For a particle with spin $1/2$, there are two possible position spots, corresponding to the two discrete results $\pm \hbar/2$ for the spin component along the direction of the magnetic field. **b**, Stern and Gerlach filter. One spot is blocked, so that particles with a spin component $+\hbar/2$ are transmitted. If the particle is transmitted once, it is prepared in the state $|+\hbar/2\rangle$, and it will always be transmitted if the measurement is repeated. This Stern and Gerlach filter is a perfect ‘quantum state preparation’ (QSP) device, but it is not a QND device because the particle would be ‘demolished’ if any external perturbation should change the spin state between two successive measurements. **c**, Simplified QND measurement of the spin. The measured particle interacts with a ‘meter’ particle carrying a magnetic moment, so that an entangled spin state of the two particles is created. The spin of the measured particle is deduced from a destructive measurement of the spin of the meter particle. This scheme can be repeated again and again, always yielding two possible results, without any demolition.

are currently performed: the measured system is coupled to another ‘meter’ system through an appropriate (conservative) interaction. This coupling builds up quantum correlations between the two systems, which remain after they have separated⁸. The quantum state after this interaction is usually in a particular quantum superposition of the measured and meter systems, called an entangled state. The correlations associated with this entangled state can give rise to the so-called Einstein–Podolsky–Rosen paradox^{28,29}, meaning that a precise measurement performed on one of the systems allows for an automatic knowledge of the state of the other. Then, the quantum state of the measured particle is determined from a direct destructive measurement on the meter particle. In such an indirect measurement, the QSP property amounts to deducing the output state of the measured particle from the spin value obtained for the meter particle. Such conditional knowledge about the outgoing measured system will play an essential role in what follows.

Once the meter particle has been detected, the result of the measurement is known, and other non-commuting observables of the measured particle (for example, its spin components along the other axis) are thus randomized. This is nothing but the back action of the measurement imposed by the Heisenberg inequality. The main point in the QND strategy is that this randomization has no unwanted effects, because it is not back-coupled onto the measured QND variable, which is fully preserved in an ideal QND measurement.

QND measurements in quantum optics

Quantum optics is a very favourable domain for implementing these ideas. This field developed rapidly during the 1980s, when considerable theoretical and experimental progress led to the generation of ‘squeezed’ states, making it possible to manipulate the quantum noise of light^{13–15,19,20}. In this context, QND ideas happened to provide an attractive way to control this quantum noise. The basic technique used in QND experiments is to couple a signal beam (to be measured) to a second beam, called the meter beam, through an optically nonlinear medium. The quantum variable to be measured in optical experiments is most frequently the number of photons during a given integration time. For typical parameters of laser beams, this photon number n is very large (typically 10^{10}); the discrete values of the photon numbers, which are very closely spaced, do not need to be resolved. A more relevant precision scale is actually set by the so-called ‘shot-noise’ limit, which is $\Delta n = \sqrt{n} \ll n$, as will be explained below. As the relative quantum fluctuations $\Delta n/n$ are then very small, they can be treated in a linear approximation. Very convenient tools in this case are the quadrature amplitudes of the electric field, which are continuous variables analogous to the position and momentum of a harmonic oscillator. In this linear approximation, it is simple to show that the photon number and phase fluctuations are given by (here we use standard notations in which $\delta\hat{A} = \hat{A} - \langle\hat{A}\rangle$ and $\Delta A = \langle(\delta\hat{A})^2\rangle^{1/2}$ for any operator $\hat{A}$): $\delta\hat{n} = \sqrt{n}\delta\hat{X}$, and $\delta\phi = \delta\hat{Y}/(2\sqrt{n})$, so that the familiar number-phase Heisenberg relation, $\Delta n\Delta\phi > 1/2$, can be written as $\Delta X\Delta Y > 1$, where now both $\hat{X}$ and $\hat{Y}$ are well behaved hermitian operators. The values $\Delta X = \Delta Y = 1$, yielding $\Delta n = \sqrt{n}$ and $\Delta\phi = 1/(2\sqrt{n})$, correspond to the so-called standard quantum limit (or shot-noise limit). Therefore, the measured quantum variable in the following will be the observable quantity $\hat{X}$.

An important step in the evolution of QND measurements in optics was the introduction in the early 1990s of a set of quantitative criteria for evaluating the information-processing ability of a given measurement scheme^{30–32}. These criteria clarified the objective to be reached, and defined a clear standard quantum limit for QND measurements^{21,33}. As they also connect our general discussion above to quantum optics experiments, we will present them briefly here.

The main idea of these criteria is simply that each of the three

requirements quoted above can be quantified using a 'measurement error', ΔX , which can be compared to the expected resolution. For instance, the QSP criterion, which states that after the measurement the system should be left in a well defined eigenstate, can be evaluated by looking at the conditional dispersion $\Delta X_{s|m}$ left in the system observable after the measurement has been made. In the case of discrete eigenvalues, $\Delta X_{s|m}$ should be smaller than the separation between two successive eigenvalues. In the case of a quasi-continuous spectrum, the expected resolution is more arbitrary, but for laser beams a very natural limit is set by the shot noise, and the required condition can be written simply as $\Delta X_{s|m} < 1$, where a value of unity corresponds to the shot-noise limit. Similarly, there is a criterion concerning the measurement efficiency which can be characterized by a quantity ΔX_m , which in the terminology of signal processing can be seen as a 'noise referred to the input' of the measurement device. Finally, the non-demolition criterion states whether or not the measured observable has been disturbed by the measurement. It can also be given as an error ΔX_s , which corresponds to the extra dispersion on the initial value due to the measurement. An efficient measurement should seek to minimize both ΔX_m and ΔX_s , and a quantitative condition, which defines the standard quantum limit for QND measurements, can be expressed by the inequality $\Delta X_s \Delta X_m < 1$ (ref. 33).

In the QND literature, it has become usual to define $T_m = 1/(1 + \Delta X_m^2)$ and $T_s = 1/(1 + \Delta X_s^2)$. These quantities are related to the transfer of signal-to-quantum-noise ratio from the input signal to the meter and the signal outputs respectively. It can also be seen that $1/T_s$, $1/T_m$ correspond to the 'noise figures' of the device, where the reference is given by the shot-noise level³³. Using these definitions, the condition $\Delta X_s \Delta X_m < 1$ can be written simply as $T_s + T_m > 1$. The joint inequalities $T_s + T_m > 1$ and $\Delta X_{s|m}^2 < 1$ are now generally accepted as necessary conditions for QND operation³³.

QND measurements with cross-phase modulation

The optical Kerr effect, frequently encountered in nonlinear optics, results from an intensity-dependent refractive index. The cross-Kerr effect corresponds to a situation in which the refractive index of a beam is modified by the intensity of a second one. In that case, the

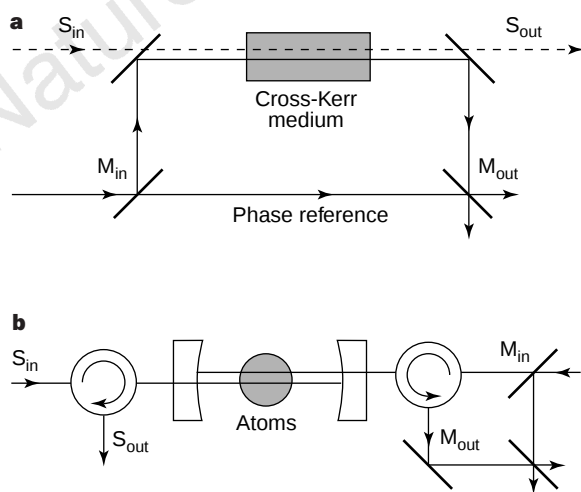


Figure 2 Measurement scheme via cross-Kerr effect. **a**, Schematic set-up for the QND measurement of the intensity of the signal beam using the cross-Kerr effect. The intensity of the signal beam alters the refractive index experienced by the meter beam. This results in a phase shift on the meter beam that is detected by placing the nonlinear medium in one arm of an interferometer, the other arm being used as a phase reference. **b**, Experimental implementation using atoms in an optical cavity^{21,22,26,42}. Rotating arrows indicate optical circulators. The quantities $S_{in,out}$, $M_{in,out}$ are, respectively, the input and output of the signal and meter beams.

refractive indices $n^{(m,s)}$ of the two beams labelled 'm' and 's' are given by $n^{(m)} = n_0^{(m)} + n_2^{(m)} I^{(s)}$, and $n^{(s)} = n_0^{(s)} + n_2^{(s)} I^{(m)}$, where $n_0^{(m,s)}$ is the linear index, $n_2^{(m,s)}$ is the nonlinear Kerr index, and $I^{(m,s)}$ is the intensity of beam m or s. This type of nonlinearity has all the ingredients for establishing quantum correlations between a signal and a meter beam, and therefore to implement intensity QND measurements^{16,18}. Using the quadrature components defined above, the input–output transformation for quantum fluctuations is:

$$\delta X_{out}^{(s)} = \delta X_{in}^{(s)} \quad \delta Y_{out}^{(s)} = \delta Y_{in}^{(s)} - g \delta X_{in}^{(m)} \quad (1)$$

$$\delta X_{out}^{(m)} = \delta X_{in}^{(m)} \quad \delta Y_{out}^{(m)} = \delta Y_{in}^{(m)} - g \delta X_{in}^{(s)} \quad (2)$$

where g is the nonlinear cross-gain, given by $g = 2\sqrt{\Phi_s \Phi_m}$. The quantities $\Phi_{s,m}$ are the nonlinear phase shifts of beam s or m, and are given by $\Phi_{s,m} = k_{s,m} l n_2^{(s,m)} I_{m,s}$, where l is the length of the nonlinear medium, and $k_{s,m}$ the wavevector of beam s or m. The important feature of this transformation is that the phase quadrature $\delta Y_{out}^{(m)}$ of the meter picks up information about the input amplitude quadrature of the signal $\delta X_{in}^{(s)}$, while the latter is left unchanged. An interferometric set up allows the direct detection of the phase of the meter beam (Fig. 2). This scheme is therefore an implementation of a non-destructive indirect measurement of the signal intensity at the quantum level, namely a QND measurement.

The QND criteria introduced earlier can easily be calculated in this ideal cross-Kerr situation; for simplicity we will assume that the input beams are shot-noise limited. As the signal intensity is unaffected by the interaction, $T_s = 1$. The quality of the measurement is evaluated by $T_m = g^2/(1 + g^2)$, and the conditional variance is $\Delta X_{s|m}^2 = 1/(1 + g^2)$. Perfect QND measurements will therefore be obtained as g becomes arbitrarily large (Fig. 3). We note that equation (2) is the useful interaction for the measurement, whereas equation (1) describes intensity noise of the meter beam being fed back onto the phase of the signal beam. This is simply the unavoidable measurement back action, which yields a signal phase

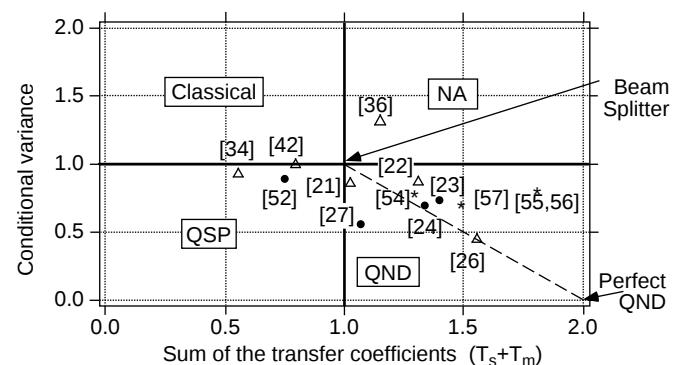


Figure 3 The different properties used to characterize a measurement device. The conditional variance, $\Delta X_{s|m}^2$, is plotted against the sum of the two transfer coefficients $T_s + T_m$. A beamsplitter lies at the point $T_s + T_m = 1$, $\Delta X_{s|m}^2 = 1$, whereas the perfect QND device would be located at the point $\Delta X_{s|m}^2 = 0$, $T_s + T_m = 2$. This diagram is divided into four regions. A system with $\Delta X_{s|m}^2 > 1$ and $T_s + T_m < 1$ has no quantum property and is therefore in the classical region. When $\Delta X_{s|m}^2 < 1$ and $T_s + T_m < 1$, the conditional variance is in the quantum domain, and the system fulfills the quantum state preparation (QSP) property. Systems located in the $\Delta X_{s|m}^2 > 1$ and $T_s + T_m > 1$ area do not establish quantum correlations between their two outputs; an example of this kind of system is a noiseless amplifier (NA) followed by a beamsplitter. The last region, $\Delta X_{s|m}^2 < 1$ and $T_s + T_m > 1$, is the only one where the QND label is deserved. The dashed line shows the behaviour predicted by equations (1, 2) for an ideal Kerr medium, as the nonlinear gain g is increased. The different experimental points are labelled with the number of the reference to which they correspond. Open triangles and filled circles represent QND experiments using third- and second-order nonlinearities, respectively. The asterisks represent 'quantum-repeater' schemes where the signal is amplified.

noise $\Delta Y_{\text{out}}^{(s)} = 1 + g^2$. As $\Delta Y_{\text{out}}^{(s)} = \Delta Y_{\text{slm}}$, then $\Delta X_{\text{slm}} \Delta Y_{\text{slm}} = 1$, which is just the Heisenberg relation for X and Y after the measurement. The minimum value of unity for the uncertainty product is obtained here because the input–output transformation (equations (1, 2)) is unitary.

The challenge now left to experimentalists is to find a real nonlinear medium that is as close as possible to an ideal cross-Kerr medium with a large cross-gain. In particular, a successful material should minimize the various perturbing effects that occur often in practice, such as absorption losses, extra noise sources, or unwanted nonlinearities. Historically, the first candidates for experimental realization of QND measurements were optical fibres^{34,35}. Although the third-order nonlinearity of silica is not very large, optical fibres can propagate light over long distances without significant losses. Moreover the possibility of operation in the soliton regime is very attractive in this context^{36–38}. The existence of other nonlinear effects and of parasitic noise sources, however, prevented these experiments from fulfilling the QND condition (Fig. 3).

Alternatively, it is well known that Kerr-type nonlinearity in resonant systems can be extremely large^{39,40}. However, these resonant nonlinear effects are often associated with very large spontaneous emission noise. For the cross-Kerr nonlinearity required in QND measurements, three-level systems are well suited. Moreover, it turns out that the degrees of freedom offered by this configuration allow very efficient control of spontaneous emission, by taking advantage of the possibility of having it almost suppressed by destructive quantum interference effects⁴¹.

In experiments with atomic vapours, the nonlinearity is further enhanced by putting the atomic medium between two mirrors forming a Fabry–Perot cavity, which magnifies the light intensity coupled to the atoms. The first QND experiments with atomic systems were performed with a beam of sodium atoms using a three-level cascade scheme⁴². This method provided the first results to fulfil the QND criteria^{21,22} (Fig. 3). However, various limitations were encountered, mostly due to residual Doppler broadening of the transitions, and to atom-number fluctuations. Laser-cooled atoms offer a convenient way to circumvent these problems. It has become possible in the past decade to trap and cool neutral atoms to temperatures well below the Doppler limit (see refs 43, 44 for reviews). The main concern when doing quantum noise experiments with this type of medium⁴⁵ is to control the mutual influence between the trapping lasers and the two beams used for the QND measurement, but this has now been achieved²⁶.

QND experiments using resonant nonlinearities in atoms have thus led to the best efficiencies so far, according to the set of QND criteria. The latest version of these experiments²⁶, employing a cloud of cold trapped atoms, currently holds the world record for QND efficiency (Fig. 3). This remarkable application of cold atoms yielded almost ‘noiseless’ optical nonlinearities, as high as nine orders of magnitude larger than in silica.

Using optical parametric amplifiers for QND

The Kerr effect results from a third-order optical nonlinearity. It is less obvious that QND effects can also be obtained from second-order nonlinearities, which are more amenable to effects like frequency doubling, twin-beam generation^{46,47}, or parametric amplification of a signal beam at frequency ω using a pump beam at frequency 2ω . Some kind of amplification process is useful for QND, because it allows the meter beam to carry out an amplified copy of the signal beam. Parametric processes are able to amplify one quadrature component of an incoming beam, while preserving its signal-to-quantum-noise ratio, with T_s close to unity. This effect is often referred to as noiseless amplification, because conventional (laser-type) optical amplifiers yield $T_s = 0.5$ for large gains, which means a 3-dB loss in signal-to-quantum-noise ratio^{48–50}. Depending on the relative phase between the pump and the signal beam,

second-order nonlinearities can also perform parametric ‘deamplification’, which when applied to a shot-noise limited beam leads to squeezed states of light^{13–15,19,20}. When impinging on the usually unused input port of a beamsplitter, squeezed states have the property of reducing the partition noise of the light which enters the ordinary input port⁵¹. The two output beams then become quantum correlated, and fulfil the condition $\Delta X_{\text{slm}}^2 < 1$. The QND scheme shown in Fig. 4 combines both effects in a single nonlinear crystal: noiseless amplification to create a copy of the signal, and noiseless splitting to separate the output signal and meter beams. This is achieved by first separating the incoming beam into two parts, one being amplified (gain G) and the other deamplified (gain $1/G$). For values of the reflection and transmission coefficients r and t leading to an overall unity gain for the signal beam, it can be shown that the input–output transformation has a form similar to equations (1, 2) by exchanging $X^{(m)}$ and $Y^{(m)}$ and taking $g = G - 1/G$. This shows that the QND variable $X_{\text{in}}^{(s)}$ is exactly recovered as $X_{\text{out}}^{(s)}$ in one output beam, while the output meter is taken from the other output beam.

The first attempt to obtain QND by this elegant method was reported by La Porta *et al.*⁵² in a pioneering paper in 1989. In practice, the scheme is implemented using type-II phase-matched crystals (for example, KTP), which can provide gain G for one polarization and $1/G$ for the orthogonal one at the same time. The two beamsplitters are simply two polarizing cubes, whose reflection and transmission coefficients are adjusted by a half-wavelength plate. However, optical losses and a low parametric gain prevented the full achievement of the QND criteria. Improved results were obtained in 1994, using a continuous-wave optical parametric amplifier (OPA) in a cavity configuration designed to enhance the parametric gain²³.

The simplicity and efficiency of this scheme allowed the central issue of repeated QND measurements to be addressed. As explained above, the main idea in QND strategy is to monitor a single observable that can be measured many times with the same result, identical to the first precise result if no external perturbation is applied. This necessity to prepare and then measure again has been

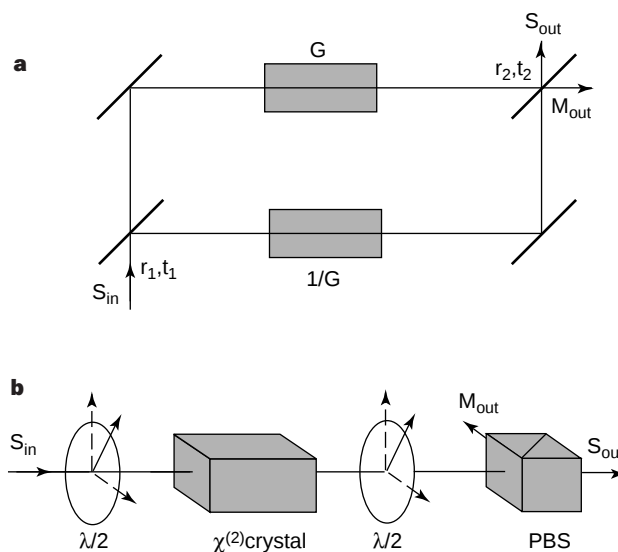


Figure 4 Measurement scheme using pulsed optical parametric amplifier with a type-II phase-matched $\chi^{(2)}$ crystal. For simplicity, the pump beams at frequency 2ω are not represented. **a**, Set-up. The input signal is amplified in one arm of the interferometer and deamplified in the other. **b**, In practical implementations^{23,24,52,54,57}, the two arms of the interferometer are two orthogonal linear polarizations. The reflection and transmission coefficients, (r_1, t_1) and (r_2, t_2) , are adjusted by rotating two half-wavelength plates, and the output signal and meter beams are obtained from a polarization beamsplitter, PBS.

pointed out since the first proposals of QND measurements, but its demonstration required the building of two independent QND setups to be operated in series, achieved in 1995²⁴. Two QND devices were implemented in a travelling-wave-pulsed OPA configuration. Both devices represented the current state of the art at that time. They were operated in series, with the signal output of the first entering the second in order to obtain a second measurement channel. The conditional variances of the signal, given the first or the second measurement, were clearly in the quantum domain. At the same time, the sum of signal transfer coefficient plus any one of the measurement transfer coefficients largely exceeded the classical limit (Fig. 3). There was a strong correlation between the quantum fluctuations of the two meter outputs (30%), demonstrating that two successive and independent measurements on the same quantum system are strongly correlated. This sequential procedure was therefore the first direct demonstration of the non-demolition property of a QND scheme.

A more recent series of experiments, performed in the continuous-wave regime, also achieved repeated QND measurement. This demonstration used an OPA as a first QND device, and a 50/50 beamsplitter assisted by squeezed light as a second one. Repeated measurements were made with stable operation for more than 10 hours²⁷.

Quantum optical repeaters

The relative simplicity of the OPA scheme in the travelling-wave configuration allowed the long-awaited goal of repeated QND measurements to be met for the first time. But while on the one hand probing the fundamentals of quantum mechanics, researchers also began to explore the consequences of quantum optics in the applied domain. A possible use of QND-type detection is related to optical communications: repeated QND measurements can be seen as a noiseless distribution of information encoded in an optical beam. Several taps disposed in series on an optical line could then extract information without any degradation for downstream users. But what about the effect of optical losses, which are inevitable as far as optical communication applications are concerned? For n QND taps separated by propagation sections with losses α , the transfer coefficient of the input signal-to-quantum-noise ratio becomes $T_s = (1 - \alpha)^n$, which rapidly tends to zero as n or/and α grow.

The fact that each QND tap keeps the signal unchanged is not good when losses occur between two successive measurements. A better way to proceed is to adjust the tapping device so that it provides (noiseless) gain, which 'precompensates' the losses that occur down the line⁵³. It turns out that this can easily be done in such a way that the QND criteria are still fulfilled. Such an amplifying 'quantum repeater' could be useful, because fulfilment of the QND criteria guarantees the possibility of repeated read-out of information encoded on an optical beam, while preserving its initial signal-to-quantum-noise ratio.

With the minor modifications to the QND configuration already described, optical parametric amplification can meet these requirements. For this purpose, the reflectivities (r_1 and r_2) are set in order to produce two magnified copies of the input signal, using the twin photon-generation property together with the noiseless amplification. By considering one of the outputs as the signal channel and the other as the measurement channel, the transfer coefficient of each output is $T_{s,m} = 1/(1 + 1/G^2)$, G being the noiseless gain. Clearly, $T_s + T_m$ approaches the maximum value of 2 for high enough gains. Furthermore, the twin outputs are quantum correlated and the conditional variance of the signal, given the measurement, goes to zero when G is increased. Such an amplifying quantum optical tap was experimentally demonstrated in 1993 (ref. 54). In parallel to this all-optical device, an opto-electronic quantum tap was also realized, based on the regeneration of the optical signal by the successive use of a high-quantum-efficiency photodetector and a light-emitting diode^{55,56}. A further step was the demonstration of an elementary

quantum optical 'bus', robust against downstream losses⁵⁷, achieved by using two identical amplifying quantum optical taps, as described before and implemented in a series configuration, after introduction of 30% losses between the two taps. The initial signal-to-quantum-noise ratio was preserved, indicating that quantum optics is able to process optical signals in the presence of losses, provided that suitable procedures are designed.

Perspectives

Originating from discussions about the fundamental quantum limits of precision experiments, QND measurements in optics have explored the ultimate limitations to the non-destructive extraction of information encoded in a laser beam. Various ways to overcome the standard quantum limit have been demonstrated in second- or third-order nonlinear interactions. The concept of QND measurement, which looked very abstract in the 1970s, has finally been realized experimentally, offering new possibilities for ultra-high-sensitivity measurements in optics.

Moreover, amplifying systems can also fulfil the set of quantitative criteria developed to characterize QND measurements, allowing robust quantum optical tapping with sub-shot-noise sensitivity. Present attempts to increase the efficiency and compactness of QND devices involve the use of new, highly nonlinear media, such as periodically poled crystals or fibres^{58,59} and microcavities. Applications of such devices are suitable whenever ultra-low-noise processing of optical signals is required.

Most of the initial objectives of optical QND measurements have now been achieved, and the domain has come to a point at which perspectives for development can be better appreciated. Present and future QND research will range from improving the reliability of QND techniques through the use of new, more compact, highly nonlinear materials, to solving technical problems, such as attaining single-photon resolution in the detection of small photon numbers⁶⁰. In the optical domain, to be able to watch individual photons fly by will be the ultimate goal of any QND measurement, although it should be emphasized that single-photon QND experiments require extremely strong coupling between the signal and meter systems. Even though they will be difficult to obtain, such ultra-high nonlinearities are within the reach of the techniques of cavity quantum electrodynamics^{61,62}.

Finally, the very techniques that should yield QND measurements for individual photons may also generate other interesting effects—for instance, a Kerr-type QND effect at the single photon level would indicate that one photon can significantly shift the phase of another light beam. But such a coupling is also the key to building a quantum logical gate, which would open the way to quantum logic and quantum computing⁶³. So the development of new experimental techniques, which yield a strong coupling between two quantum systems and which originate from work on QND measurements, now promises many other opportunities. □

P. Grangier and J.-P. Poizat are at the Laboratoire Charles Fabry de l'Institut d'Optique, CNRS (URA 14), BP 147, F-91403 Orsay Cedex, France; J. A. Levenson is at the France Telecom/CNET/DTL/Laboratoire de Bagnex, CNRS (URA 250), BP 107, F-92225 Bagneux Cedex, France.

1. Wheeler, J. A. & Zurek, W. H. (eds) *Quantum Theory and Measurement* (Princeton Univ. Press, Princeton, NJ, 1983).
2. Braginsky, V. B. & Vorontsov, Y. I. Quantum-mechanical limitations in macroscopic experiments and modern experimental technique. *Usp. Fiz. Nauk* **114**, 41–53 (1974) [*Sov. Phys. Usp.* **17**, 644–650 (1975)].
3. Braginsky, V. B., Vorontsov, Y. I. & Khalili, F. Y. Quantum singularities of a ponderomotive meter of electromagnetic energy. *Zh. Eksp. Teor. Fiz.* **73**, 1340–1343 (1977) [*Sov. Phys. JETP* **46**, 705–706 (1977)].
4. Thorne, K. S., Drever, R. W. P., Caves, C. M., Zimmermann, M. & Sandberg, V. D. Quantum nondemolition measurements of harmonic oscillators. *Phys. Rev. Lett.* **40**, 667–671 (1978).
5. Unruh, W. G. Quantum nondemolition and gravity-wave detection. *Phys. Rev. B* **19**, 2888–2896 (1979).
6. Caves, C. M., Thorne, K. S., Drever, R. W. P., Sandberg, V. D. & Zimmermann, M. On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principles. *Rev. Mod. Phys.* **52**, 341–392 (1980).
7. Braginsky, V. B., Vorontsov, Y. I. & Thorne, K. S. Quantum nondemolition measurements. *Science* **209**, 547–557 (1980).

8. Braginsky, V. B. & Khalili, F. Y. *Quantum Measurement* (ed. Thorne, K. S.) (Cambridge Univ. Press, Cambridge, 1992).
9. Brillet, A., Damour, T. & Tournenc, P. Introduction to gravitational wave research. *Ann. Phys. Fr.* **10**, 201–218 (1985).
10. Brillat, A. Interferometric gravitational wave antennae. *Ann. Phys. Fr.* **10**, 219–226 (1985).
11. Weber, J. Gravitational-wave-detector events. *Phys. Rev. Lett.* **20**, 1307–1308 (1968).
12. Bocko, M. F. & Onofrio, R. On the measurement of a weak classical force coupled to a harmonic oscillator: experimental progress. *Rev. Mod. Phys.* **68**, 755–799 (1996).
13. Yuen, H. P. Two-photon coherent states of the radiation field. *Phys. Rev. A* **13**, 2226–2243 (1976).
14. Caves, C. M. Quantum-mechanical noise in an interferometer. *Phys. Rev. D* **23**, 1693–1708 (1981).
15. Walls, D. F. Squeezed states of light. *Nature* **306**, 141–146 (1983).
16. Milburn, G. J. & Walls, D. F. Quantum nondemolition measurements via quadratic coupling. *Phys. Rev. A* **28**, 2065–2070 (1983).
17. Yurke, B. Optical back-action-evading amplifiers. *J. Opt. Soc. Am. B* **2**, 732–738 (1985).
18. Imoto, N., Haus, H. A. & Yamamoto, Y. Quantum nondemolition measurement of the photon number via the optical Kerr effect. *Phys. Rev. A* **32**, 2287–2292 (1985).
19. Squeezed light. *J. Opt. Soc. Am. B* **4** (suppl.), 1449–1741 (1987).
20. Squeezed light. *J. Mod. Opt.* **34**, 709–1020 (1987).
21. Roch, J. F., Roger, G., Grangier, P., Courty, J. M. & Reynaud, S. Quantum non-demolition measurements in optics: a review and some recent experimental results. *Applied Phys. B* **55**, 291–297 (1992).
22. Poizat, J. P. & Grangier, P. Experimental realization of a quantum optical tap. *Phys. Rev. Lett.* **70**, 271–274 (1993).
23. Pereira, S. F., Ou, Z. Y. & Kimble, H. J. Backaction evading measurements for quantum nondemolition detection and quantum optical tapping. *Phys. Rev. Lett.* **72**, 214–217 (1994).
24. Bencheikh, K., Levenson, J. A., Grangier, P. & Lopez, O. Quantum nondemolition demonstration via repeated backaction evading measurements. *Phys. Rev. Lett.* **75**, 3422–3425 (1995).
25. Quantum nondemolition measurements. *Appl. Phys. B* **64** (suppl.), 123–272 (1997).
26. Roch, J.-F. *et al.* Quantum nondemolition measurements using cold atoms. *Phys. Rev. Lett.* **78**, 634–637 (1997).
27. Bruckmeier, R., Hansen, H. & Schiller, S. Repeated quantum nondemolition measurements of continuous optical waves. *Phys. Rev. Lett.* **79**, 1463–1466 (1997).
28. Einstein, A., Podolsky, B. & Rosen, N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777–780 (1935).
29. Bohr, N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **48**, 696–702 (1935).
30. Imoto, N. & Saito, S. Quantum nondemolition measurement of photon number in a lossy optical Kerr medium. *Phys. Rev. A* **39**, 675–682 (1989).
31. Holland, M. J., Collett, M. J., Walls, D. F. & Levenson, M. D. Nonideal quantum nondemolition measurements. *Phys. Rev. A* **42**, 2995–3005 (1990).
32. Grangier, P., Courty, J. M. & Reynaud, S. Characterization of nonideal quantum nondemolition measurements. *Opt. Commun.* **89**, 99–106 (1992).
33. Poizat, J. P., Roch, J. F. & Grangier, P. Characterization of quantum nondemolition measurements in optics. *Ann. Phys. Fr.* **19**, 265–297 (1994).
34. Levenson, M. D., Shelby, R. M., Reid, M. & Walls, D. F. Quantum nondemolition detection of optical quadrature amplitudes. *Phys. Rev. Lett.* **57**, 2473–2476 (1986).
35. Bachor, H. A., Levenson, M. D., Walls, D. F., Perlmutter, S. H. & Shelby, R. M. Quantum nondemolition measurements in an optical-fiber ring resonator. *Phys. Rev. A* **38**, 180–190 (1988).
36. Friberg, S. R., Machida, S. & Yamamoto, Y. Quantum nondemolition measurement of the photon number of an optical soliton. *Phys. Rev. Lett.* **69**, 3165–3168 (1992).
37. Drummond, P. D., Shelby, R. M., Friberg, S. R. & Yamamoto, Y. Quantum solitons in optical fibres. *Nature* **365**, 307–313 (1993).
38. Friberg, S. R. Quantum nondemolition: probing the mystery of quantum mechanics. *Science* **278**, 1088–1089 (1997).
39. Grangier, P., Roch, J. F. & Reynaud, S. Quantum correlations and non-demolition measurements using two-photon non-linearities in optical cavities. *Opt. Commun.* **72**, 387–392 (1989).
40. Levenson, M. D. *et al.* Cross-quadrature modulation with the Raman-induced Kerr effect. *Phys. Rev. A* **44**, 2023–2035 (1991).
41. Gheri, K. M., Grangier, P., Poizat, J. P. & Walls, D. F. Quantum nondemolition measurements using ghost transitions. *Phys. Rev. A* **46**, 4276–4285 (1992).
42. Grangier, P., Roch, J. F. & Roger, G. Observation of backaction-evading measurement of an optical intensity in a three-level atomic nonlinear system. *Phys. Rev. Lett.* **66**, 1418–1421 (1991).
43. Cohen-Tannoudji, C. in *Fundamental Systems in Quantum Optics* (eds Dalibard, J., Raymond, J.-M. & Zinn-Justin, J.) 1–164 (North Holland, Amsterdam, 1992).
44. Phillips, W. D. in *Fundamental Systems in Quantum Optics* (eds Dalibard, J., Raymond, J.-M. & Zinn-Justin, J.) 165–210 (North Holland, Amsterdam, 1992).
45. Lambrecht, A., Coudreau, T., Steinberg, A. M. & Giacobino, E. Squeezing with cold atoms. *Europhys. Lett.* **36**, 93–98 (1996).
46. Heidmann, A. *et al.* Observation of quantum noise reduction on twin laser beams. *Phys. Rev. Lett.* **59**, 2555–2557 (1987).
47. Aytur, O. & Kumar, P. Pulsed twin beams of light. *Phys. Rev. Lett.* **65**, 1551–1554 (1990).
48. Caves, C. M. Quantum limits on noise in linear amplifiers. *Phys. Rev. D* **26**, 1817–1839 (1982).
49. Ou, Z. Y., Pereira, S. F. & Kimble, H. J. Quantum noise reduction in optical amplification. *Phys. Rev. Lett.* **70**, 3239–3242 (1993).
50. Levenson, J. A., Abram, I., Riviera, T. & Grangier, P. Reduction of quantum noise in optical parametric amplification. *J. Opt. Soc. Am. B* **10**, 2233–2238 (1993).
51. Shapiro, J. H. Optical waveguide tap with infinitesimal insertion loss. *Opt. Lett.* **5**, 351–353 (1980).
52. La Porta, A., Slusher, R. E. & Yurke, B. Back-action evading measurements of an optical field using parametric down conversion. *Phys. Rev. Lett.* **62**, 28–31 (1989).
53. Yuen, H. P. Generation, detection, and application of high-intensity photon-number-eigenstates fields. *Phys. Rev. Lett.* **56**, 2176–2179 (1986).
54. Levenson, J. A. *et al.* Quantum optical cloning amplifier. *Phys. Rev. Lett.* **70**, 267–270 (1993).
55. Goobar, E., Karlsson, A. & Björk, G. Experimental realization of a semiconductor photon number amplifier and a quantum optical tap. *Phys. Rev. Lett.* **71**, 2002–2005 (1993).
56. Roch, J.-F., Poizat, J.-P. & Grangier, P. Subshotnoise manipulation of light using semiconductor emitters and receivers. *Phys. Rev. Lett.* **71**, 2006–2009 (1993).
57. Bencheikh, K., Simonneau, C. & Levenson, J. A. Cascaded amplifying quantum optical taps: a robust noiseless optical bus. *Phys. Rev. Lett.* **78**, 34–37 (1997).
58. Fejer, M. M., Magel, G. A., Jundt, D. H. & Byer, R. Quasi-phase-matched second harmonic generation: tuning and tolerances. *IEEE J. Quant. Electr.* **28**, 2631–2654 (1992).
59. Løvering, D. J., Levenson, J. A., Vidakovic, P., Webjörn, J. & Russel, P. St. J. Noiseless optical amplification using quasi-phase-matched bulk lithium niobate. *Opt. Lett.* **21**, 1439–1441 (1996).
60. Brune, M., Haroche, S., Lefevre, V., Raimond, J. M. & Zagury, N. Quantum nondemolition measurement of small photon numbers by Rydberg-atom phase-sensitive detection. *Phys. Rev. Lett.* **65**, 976–979 (1990).
61. Brune, M. *et al.* Quantum Rabi oscillation: a direct test of field quantization in a cavity. *Phys. Rev. Lett.* **76**, 1800–1803 (1996).
62. Turchette, Q. A., Hood, C. J., Lange, W., Mabuchi, H. & Kimble, H. J. Measurement of conditional phase shifts for quantum logic. *Phys. Rev. Lett.* **75**, 4710–4713 (1995).
63. Ekert, A. & Jozsa, R. Quantum computation and Shor's factoring algorithm. *Rev. Mod. Phys.* **68**, 733–753 (1996).

Acknowledgements. We acknowledge decisive contributions from K. Bencheikh and J.-F. Roch in our experiments. This work was supported in part by the European ESPRIT program.

Correspondence should be addressed to P.G. (e-mail: philippe.grangier@iota.u-psud.fr).

Defining the functions of *trans*-SNARE pairs

Christian Ungermann*, Ken Sato*† & William Wickner*

* Dartmouth Medical School, Department of Biochemistry, 7200 Vail, Hanover, New Hampshire 03755, USA

The homotypic fusion of yeast vacuoles includes a 'docking' step, which we show here to consist of two sequential reactions: a reversible 'tethering' mediated by the GTPase Ypt7, and 'SNARE pairing', in which SNARE proteins from opposite membranes form a complex in *trans*. The function of this *trans*-SNARE complex must be transient, as the complex can be disassembled by excess Sec18 in the presence of Sec17 and ATP without influencing the fusion rate. These data indicate that SNARE pairing may transiently signal to downstream factors, leading to fusion.

Regulated membrane fusion is essential for the formation and maintenance of subcellular compartments. From yeast to neurons, fusion requires ATPases (Sec18/N-ethylmaleimide-sensitive fusion protein (NSF)/p97), accessory proteins (Sec17/soluble NSF-attachment proteins (SNAP)), integral membrane SNAP receptors (SNAREs), and GTPases Rab/Ypt¹⁻⁶. SNAREs form complexes in *cis* on the same membrane or in *trans* to promote interorganelle docking^{7,8}. SNARE complexes are disassembled by NSF/Sec18 and α -SNAP/Sec17 (refs 1–6). The mechanism by which SNAREs promote docking and the role of *trans*-SNARE complexes in subsequent membrane fusion have been intensively studied⁴⁻⁶. Docking also requires Rab/Ypt proteins and 'velcro factors'^{9,10}. For example, Ypt1 and Uso1 'tether' endoplasmic-reticulum-derived vesicles to the Golgi¹¹ before the vesicle-bound (v)-SNAREs Bet1 and Bos1 act^{11,12}. Furthermore, the initial binding of synaptic vesicles to the presynaptic plasma membrane in the nerve terminal is also independent of SNARE proteins, as synaptic vesicles remain bound despite SNARE cleavage by neurotoxins or deletion of the SNAREs syntaxin or synaptobrevin^{13,14}. Purified SNAREs reconstituted in opposing liposomes have been shown to promote an exchange of membrane lipid, which may represent bilayer fusion¹⁵.

We have studied the fusion of yeast vacuoles, the last step in the inheritance of this organelle^{16,17}. Our *in vitro* reaction occurs in obligately ordered stages of vacuole priming, docking and fusion. During priming, a vacuolar SNARE complex containing the target membrane (t)-SNARE Vam3, the SNAP-25 (synaptosome-associated protein of relative molecular mass 25,000) homologue Vam7, the v-SNARE Nyv1, Sec17/ α -SNAP, Sec18/NSF and LMA1 is disassembled by Sec17/Sec18/ATP^{3,18,19}. Docking to the target vacuole depends on priming and on Ypt7 (ref. 20). The final fusion event is sensitive to GTP- γ S, to the phosphatase inhibitor microcystin LR (MCLR) and to mastoparan^{21,22}, although the proteins that are sensitive to these compounds have not yet been identified.

We now show that vacuole tethering requires Ypt7 but not the SNAREs and that '*trans*' v-SNARE/t-SNARE pairs function transiently, leading to further steps that directly catalyse membrane fusion. In our fusion assay we use vacuoles purified from two strains. Strain BJ3505 is deficient in vacuole luminal proteases and thus accumulates catalytically inactive pro-alkaline phosphatase. Strain DKY6281 has normal vacuolar proteases but, because of gene deletion, has no alkaline phosphatase. Fusion allows the luminal protease from DKY6281 vacuoles to act upon the pro-alkaline phosphatase from BJ3505 vacuoles, yielding mature and active alkaline phosphatase which can be assayed colorimetrically^{17,22}. Proteolytic activation of the pro-

phatase occurs almost immediately after fusion, as GTP- γ S or MCLR, which would not inhibit proteases, block activation as rapidly as transfer to ice²³. Assays have also been developed for the sequential subreactions of priming^{3,22} and docking²⁰.

Docking comprises both tethering and SNARE pairing

vacuolar SNAREs are initially found in a multisubunit *cis*-SNARE complex^{3,18} which is disassembled by Sec17/Sec18/ATP, liberating

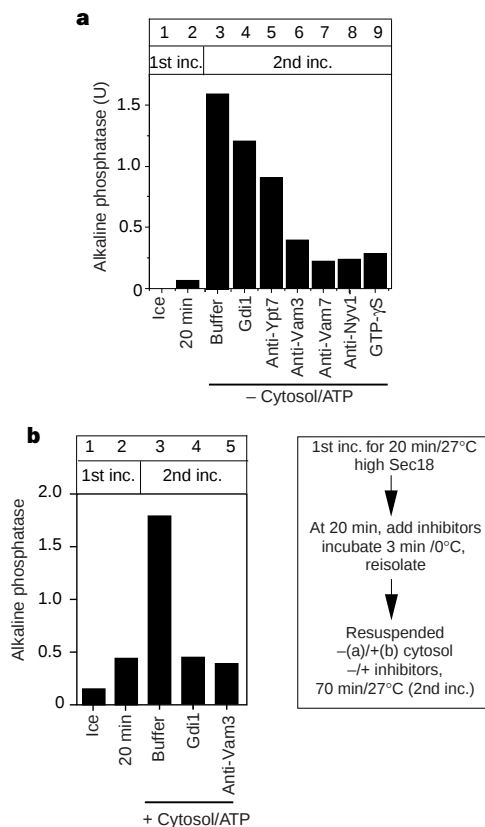


Figure 1 Resolution of tethering and SNARE function. Standard fusion reactions³ were incubated with cytosol and 3.8 μ g Sec18 for 20 min at 27 °C or were kept on ice (lane 1). After 20 min, one reaction was put on ice (lane 2); the others received inhibitors or antibodies and were placed on ice for 3 min. Reactions were diluted with 500 μ l PS buffer (10 mM Pipes/KOH, pH 6.8, 200 mM sorbitol). Vacuoles were sedimented (5 min, 10,000g, 4 °C), resuspended in the original volume of **a**, reaction buffer or **b**, reaction buffer with cytosol and ATP, mixed with the indicated inhibitors^{4,23}, incubated for 70 min at 27 °C, and put on ice. Fusion activity was measured³² and background activity (0.25 units) was subtracted.

† Present address: Division of Biological Sciences, Nagoya University, Chikusa-ku, Nagoya 464-8206, Japan.

the SNAREs for subsequent docking with SNAREs from opposing partner vacuoles to form a complex in *trans*. Both Ypt proteins and SNAREs are required for docking^{4,6,9,11}. We now show, in three independent types of experiment, that Ypt7 and SNAREs act in sequential 'tethering' and 'SNARE-pairing' stages of docking.

As Sec18 efficiently disassembles *cis*-SNARE complexes during priming, it seemed possible that excess Sec18 might disassemble *trans*-SNARE complexes before they could fulfil their functions, yet allow the Ypt7 function to proceed. To test this (Fig. 1a), we started a fusion reaction in the presence of ATP and excess Sec18. Little fusion had occurred after 20 min (Fig. 1a, lane 2), reflecting a kinetic delay in fusion caused by Sec18 (data not shown). Inhibitors were then added, the vacuoles were re-isolated to remove ATP, cytosol and the excess Sec18 and were resuspended in buffer alone or with the same inhibitors. Fusion was measured after a further 70-min incubation. Addition of Gdi1 (a guanine-nucleotide-dissociation protein (GDI), which counteracts the effects of Ypt7) or antibody against Ypt7 to the second incubation caused only a modest inhibition of the fusion reaction (Fig. 1a, lanes 4, 5), indicating that Ypt7 had completed its function on most vacuoles. Addition of either of these reagents from the start of the first reaction inhibited fusion completely²³. However, the addition of antibodies against SNAREs in the second reaction blocked almost all fusion (Fig. 1a, lanes 6–8). In incubations begun with only normal levels of Sec18, the reaction gained equivalent resistance to Gdi1 and anti-SNARE antibodies after 20 min (see below). Thus, when SNARE pairing is actively suppressed, Ypt7 function can be shown to precede SNARE function and, as shown below, support reversible vacuole tethering.

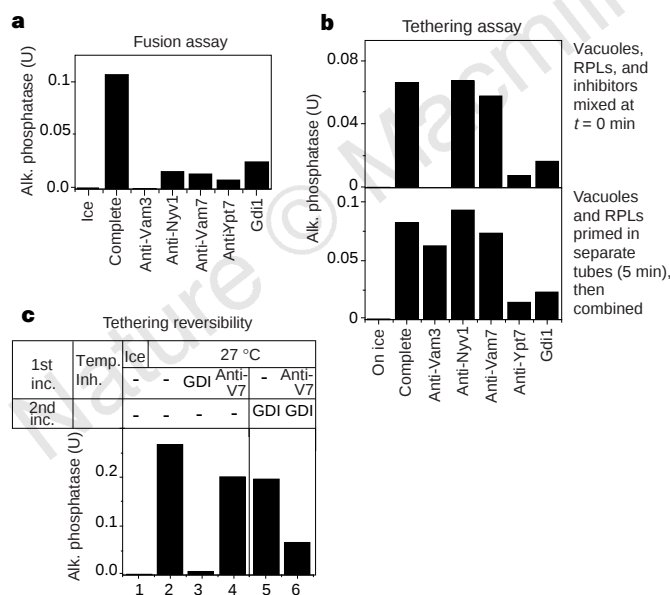


Figure 2 Tethering requires Ypt7 but not SNAREs. **a**, Proteoliposome fusion²⁴ with intact vacuoles. Fusion, as measured by alkaline phosphatase activity, is sensitive to anti-SNARE and anti-Ypt7 antibodies and Gdi1. The activity of mature phosphatase was measured after subtracting background activity (on ice; 0.05 U). **b**, Tethering assay (see Methods). Background activities of RPLs^M in reaction buffer (0.061 U) and ice control (0.110 U) were subtracted from the ALP activities. **c**, Tethering is reversible. RPLs were prepared from the wild-type strain RSY249 and incubated with MCLR, vacuoles from strain DKY6281, and (where indicated) Gdi1 (GDI) or anti-Vam7 antibody (anti-V7) for 90 min at 27 °C or on ice. Gdi1 was then added (lanes 5, 6) to complete reaction or to one that had received anti-V7 from the start to block SNARE function. The reaction was continued for 5 min at 27 °C. Tethering was assayed as in **b** (see Methods).

To further resolve the steps of docking, we developed a co-sedimentation assay. For this assay, we used reconstituted proteoliposomes (RPLs) that can fuse with intact vacuoles²⁴. This fusion is sensitive to Gdi1 and to anti-Ypt7 and anti-SNARE antibodies (Fig. 2a). The small size and low density of RPLs allows a second, independent assay of tethering: proteoliposomes were reconstituted from a mixture of mature, active alkaline phosphatase and a detergent extract of the protease-deficient BJ3505 vacuoles. These RPLs were then incubated with BJ3505 vacuoles, ATP, cytosol and the indicated inhibitors (Fig. 2b). After 45 min, we isolated vacuoles and assayed the alkaline phosphatase activity from the RPLs that had co-sedimented with vacuoles. Because we used Pep4-deficient BJ3505 vacuoles, no additional mature alkaline phosphatase would be formed upon fusion and thus the assay measured only tethering or docking. Binding of RPLs to vacuoles (Fig. 2b) was blocked by Gdi1 or anti-Ypt7 antibody, indicating that Ypt7 is crucial in the initial tethering step of the docking reaction. In contrast, antibodies to the SNAREs Vam7 or Nyv1 did not interfere with tethering. Addition of anti-Vam3 antibodies at the beginning of the incubation blocked tethering (Fig. 2b, top). However, antibodies against Vam3 (but not anti-Nyv1 antibodies) block the priming reaction³; priming is a prerequisite of docking²⁰. These results indicate that tethering may require priming. When we incubated vacuoles and RPLs in separate tubes for 5 min to allow priming and then combined them in the presence of the antibodies, anti-SNARE antibodies did not influence tethering (Fig. 2b, bottom). After 5 min of the reaction, these antibodies can still react with vacuolar SNAREs, as they still inhibit fusion³. Although we do not know

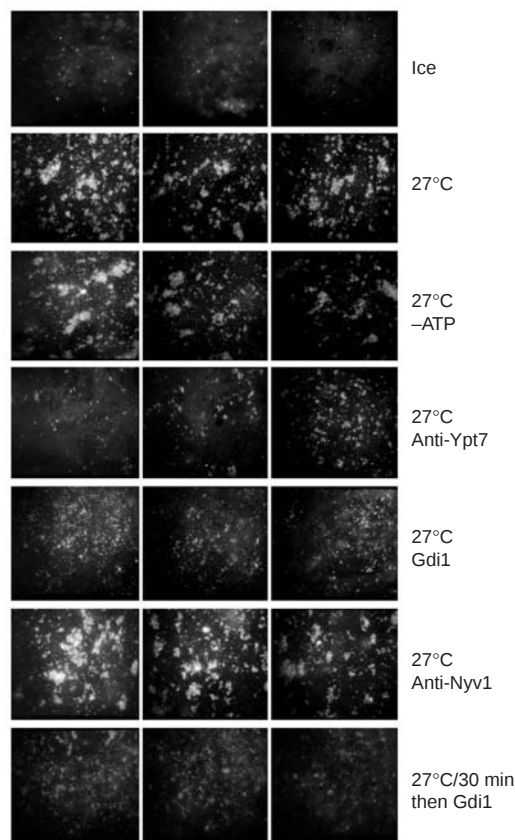


Figure 3 Vacuole tethering requires Ypt7 but not SNAREs. Vacuoles from strain DKY6281v_{am3}Δ were incubated at 27 °C for 30 min and assayed for docking²⁰. Where indicated, Gdi1 was added at 30 min and the incubation was continued for 5 min. Vacuoles were labelled with the dye FM4-64 (Molecular Probes) and random fields were photographed. The higher total fluorescence of samples with clustered vacuoles is probably due to these clusters settling onto the slide.

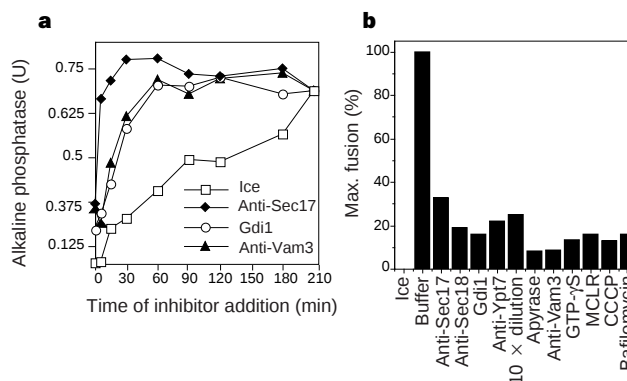


Figure 4 Fusion of vacuoles lacking t-SNAREs with those lacking v-SNAREs. **a**, Kinetics of fusion. Reactions (750 μ l) contained vacuoles from BJ3505nyv1 Δ and DKY6281vam3 Δ strains, cytosol and ATP at 27°C. Aliquots (30 μ l) were removed at indicated times, placed on ice or added to Gdi1 or antibodies against Sec17 or Vam3, incubated at 27°C for a total of 210 min, and assayed for phosphatase activity. **b**, The fusion of Δt and Δv vacuoles is sensitive to each known reaction

inhibitor. Inhibitors were added to the reaction described in **a** as follows: GTP- γ S (2 mM), MCLR (10 μ M), carbonyl cyanide-*m*-chlorophenyl hydrazone (CCCP) (10 μ M), bafilomycin (10 μ M), or apyrase (10 U ml⁻¹). Gdi1 or antibodies against Sec17, Sec18, Vam3 or Ypt7 were added at the reported inhibitory concentrations^{3,23}. Dilution was with reaction buffer containing cytosol and ATP.

how the priming signal is transmitted to the tethering reaction, our results show that tethering requires priming and depends on Ypt7 but is independent of SNARE proteins.

In a third, independent assay of tethering we used vacuoles from a *vam3* Δ strain that do not require priming³. These vacuoles cannot form v-SNARE/t-SNARE pairs but tether in a temperature-dependent and ATP-independent manner (Fig. 3, top three panels), consistent with the earlier finding that these Δt -SNARE vacuoles can fuse with Δv -SNARE vacuoles without ever having been primed³. Gdi1 or anti-Ypt7 antibodies blocked this tethering, whereas anti-Nyv1 antibodies had no effect.

Trans-SNARE pairs stabilize docking

Surprisingly, addition of Gdi1 after successful tethering led to dispersion of the vacuoles (Fig. 3, bottom panel). We therefore tested whether this reversibility of tethering is lost when SNAREs have paired. We incubated RPLs, prepared from wild-type vacuoles, with DKY6281 vacuoles in the presence of the fusion inhibitor MCLR (Fig. 2c). After 90 min, we added Gdi1 to a control reaction (Fig. 2c, lane 5) and to one that had received anti-Vam7 antibodies from the start of the initial, 90-min incubation to block SNARE function (Fig. 2c, lane 6). Strikingly, Gdi1 did not reverse the Ypt7-dependent tethering in reactions in which SNAREs could pair (Fig. 2c, lane 5) but was still capable of reversing tethering if SNARE function had been compromised (Fig. 2c, lane 6). Thus Ypt7 is required continuously for tethering contact between vacuoles until SNARE pairing makes docking irreversible. These experiments (Figs 2, 3), in which tethering was reversible, were done in the continuous presence of cytosol; tethering was resistant to the addition of Gdi1 or antibody to Ypt7 after the vacuoles had been separated from the cytosol by centrifugation and resuspension (Fig. 1a). As expected, tethering as measured by this assay was also reversed by Gdi1 when cytosol was included in the resuspension buffer (Fig. 1b), providing an assay for isolation of the cytosolic factor(s) that support the reversal of tethering.

t-SNARE/v-SNARE pairing during docking

To test directly whether SNAREs pair during docking, we prepared vacuoles from yeast strains lacking genes for either the v-SNARE Nyv1p or the t-SNARE Vam3 (ref. 25). Because of the deletion of these genes, neither mutant vacuole contained a Nyv1–Vam3 SNARE complex. These vacuoles have only half the capacity to form *trans*-SNARE pairs, and indeed the fusion of Δt (*vam3* Δ) with Δv (*nyv1* Δ) vacuoles (Fig. 4a) occurs at half the rate of wild-type

fusion. Samples received an inhibitor at various times during the incubation. Antibodies against Sec17 blocked priming when added at early times but had no effect at later times, as reported for wild-type vacuoles²³. The kinetics of inhibition by Gdi1 and anti-Vam3 antibodies, inhibitors of tethering and SNARE pairing, respectively, are not resolved, as for wild-type vacuoles (C.U. & W.W., unpublished observations). Priming (inhibited by anti-Sec17 antibodies)

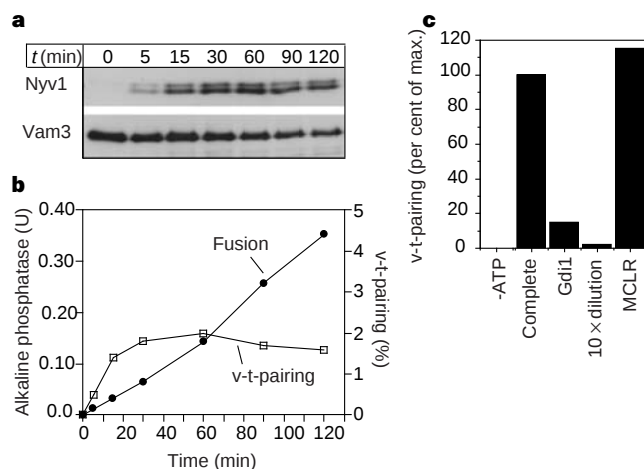


Figure 5 Formation of v-SNARE/t-SNARE pairs. **a**, **b**, Co-immunoprecipitation of Nyv1 with Vam3. Vacuoles from strains BJ3505nyv1 Δ and DKY6281vam3 Δ (100 μ g each) were incubated together at 27°C in a 750- μ l reaction with cytosol and ATP. Samples were placed on ice at the indicated times. A 30- μ l portion was used to measure ALP activity (**b**). The remainder was centrifuged (10 min, 16,000g, 4°C), suspended in 500 μ l PS buffer and sedimented. Protein complexes (**a**, **b**) were analysed by co-immunoprecipitation in 0.5% Triton X-100, phosphate-buffered saline, 2 mM EDTA, 1 $\times$ protease inhibitor cocktail²⁴ and 1 mM phenylmethylsulphonyl fluoride with protein A-immobilized antibodies to Vam3 (ref. 3). Proteins were released by 900 μ l of 0.1 M glycine, pH 2.6, and precipitated with trichloroacetate, washed with cold acetone, and analysed by SDS-PAGE and immunoblot with antibodies against Vam3 and Nyv1. The Nyv1 band (**a**) was quantified by densitometry and plotted as the percentage of total Nyv1 (**b**). **c**, Sensitivity of v-SNARE/t-SNARE pairing to docking and fusion inhibitors. Mixed *vam3* Δ and *nyv1* Δ vacuoles were incubated with cytosol, ATP and inhibitors (at concentrations described in Fig. 4b) for 40 min at 27°C, then sedimented and analysed for co-immunoprecipitation as in Fig. 4b. Maximal Nyv1 co-precipitation with Vam3 without inhibitor was set to 100%.

and docking (inhibited by anti-Vam3 antibodies and Gdi1) are rapid, and fusion *per se* is rate-limiting for the reaction. The fusion of Δt and Δv vacuoles is sensitive to each known reaction inhibitor (Fig. 4b) and is thus similar to wild-type vacuole fusion.

To assay the formation of v-SNARE/t-SNARE complexes between Δv and Δt vacuoles, we prepared detergent extracts and measured the co-immunoprecipitation of the v-SNARE Nyv1 by antibody against the t-SNARE Vam3 (Fig. 5a). Whereas the amount of v-SNARE/t-SNARE pairs was nearly maximal after 20 min, little fusion had occurred at this time (Fig. 5b), indicating that *trans*-SNARE complexes may accumulate at an intermediate reaction state, such as docking. Vacuole SNAREs are not abundant, and the engagement of a low percentage of such SNAREs in SNARE pairs may correspond to only one or two SNARE pairs per docked vacuole. Indeed, 95% of the Nyv1 in detergent extracts sedimented as an uncomplexed monomer under conditions of maximal SNARE-pair formation (data not shown). The formation of *trans*-v-SNARE/t-SNARE complexes is sensitive to Gdi1 or dilution of the reaction, but not to the irreversible late-stage inhibitor MCLR (Fig. 5c). Thus, SNAREs from separate vacuoles form v-SNARE/t-SNARE complexes in *trans* in a reaction that requires tethering but not fusion.

Dissociation of *trans*-SNARE complexes

As *cis*-SNARE complexes on vacuoles are disassembled by Sec17/Sec18 during priming¹⁸, we asked whether *trans*-SNARE complexes can also be dissociated (Fig. 6a). Isolated *vam3* Δ and *nyv1* Δ

vacuoles were incubated first with cytosol in the absence or presence of ATP for 40 min at 27 °C, and then with combinations of Sec17, Sec18 and ATP for another 10 min. Detergent extracts were assayed for the co-immunoprecipitation of Nyv1 by antibodies against Vam3. A *trans*-SNARE complex between Nyv1 and Vam3 had accumulated after 40 min in the presence of ATP and cytosol (Fig. 6a, lane 2). It was dissociated by Sec18 and ATP (Fig. 6a, lane 4) and Sec17 enhanced the dissociation (Fig. 6a, lane 5). Thus, Sec17, Sec18 and ATP can also dissociate a SNARE complex formed in *trans*. These complexes were not merely *cis* complexes that resulted from completed fusion events, as they also accumulated in the presence of the fusion inhibitor MCLR (Fig. 5c) and were still a target for Sec17/Sec18/ATP-mediated disassembly in the presence of this inhibitor (data not shown). Disassembly occurred on the vacuole, and not in detergent solution, as Nyv1–Vam3 pairs that formed at 27 °C (Fig. 6b, lane 2) were still disassembled when apyrase (which catalyses the hydrolysis of ATP) was added at the end of the incubation with Sec18 and ATP (Fig. 6b, lane 3), although disassembly did not occur when apyrase was added with Sec18 at the start of the incubation (Fig. 6b, lane 4) or when the vacuoles were incubated with ATP alone and mixed with Sec18 and apyrase only at the end of the incubation (Fig. 6b, lane 5).

As *trans*-SNARE complexes are dissociated by Sec17/Sec18, we tested whether dissociation affects fusion. Increasing the amount of Sec18 from the start of the incubation strongly reduced the steady-state level of v-SNARE/t-SNARE pairs and, although it caused kinetic delay (Fig. 1a, lane 2), it had far less effect on the overall

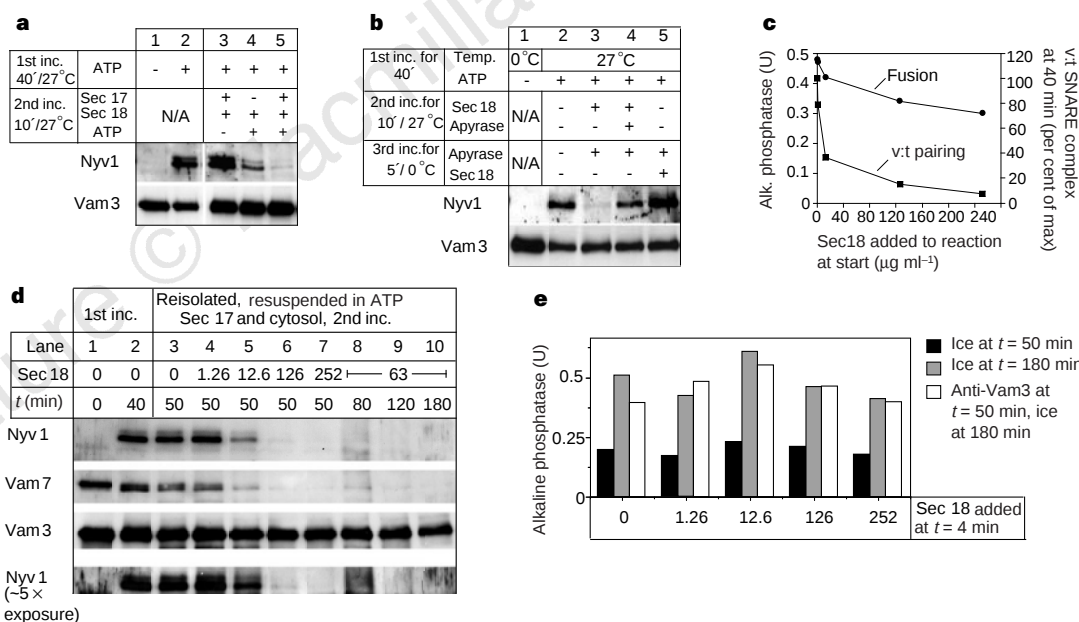


Figure 6 Dissociation of the *trans*-SNARE complex. **a**, Dissociation of the *trans* v-SNARE/t-SNARE complex by Sec18/Sec17 and ATP. Δt and Δv vacuoles (100 μg each) were incubated together with or without ATP for 40 min at 27 °C, sedimented (5 min, 16,000g, 4 °C), and either put on ice (lanes 1, 2) or resuspended with the indicated combinations of His₆-Sec18 (12.6 $\mu\text{g ml}^{-1}$) (ref. 33), His₆-Sec17 (2.7 $\mu\text{g ml}^{-1}$) and 0.5 mM Mg-ATP with an ATP-regenerating system³² (lanes 3–5). After 10 min at 27 °C, vacuoles were re-isolated and analysed as in Fig. 5a. **b**, Dissociation of the v-SNARE/t-SNARE complex requires the action of Sec18 and ATP on the intact vacuole. The first incubation and vacuole sedimentation were as in **a**. Vacuoles were then resuspended with Sec17 (2.7 $\mu\text{g ml}^{-1}$), cytosol, 0.5 mM Mg-ATP with an ATP-regenerating system and His₆-Sec18 (25 $\mu\text{g ml}^{-1}$) as shown. Apyrase (40 U ml^{-1}) was added to the premixed reaction buffer 10 min before resuspension (lane 4). After 10 min at 27 °C, vacuoles were sedimented and resuspended in PS buffer with (lanes 3–5) or without (lane 2) 5 U ml^{-1} apyrase. In lanes 3, 5, the apyrase concentration was increased to 40 U ml^{-1} , purified Sec18

premixed with 1 mM Mg-ATP was added (lane 5), and samples were left on ice for 5 min. Vacuoles were then re-isolated, washed once with PS buffer, and analysed as in Fig. 5a. **c**, Excess Sec18 affects the rate of fusion and the steady-state level of v-SNARE/t-SNARE pairing to different degrees. Fusion reactions with the indicated amounts of Sec18 were incubated as in **a** for 180 min and assayed for SNARE pairs and for fusion. **d**, **e**, Dissociation of *trans*-SNARE complexes does not influence the fusion rate. Fusion reactions were as in Fig. 5a. After 40 min, vacuoles were re-isolated (5 min, 16,000g, 4 °C) and resuspended in buffer containing ATP, cytosol and Sec17 (2.7 $\mu\text{g ml}^{-1}$). Where indicated, purified Sec18 was added at 1.26 $\mu\text{g ml}^{-1}$ (optimal for fusion), 12.6 $\mu\text{g ml}^{-1}$, 63 $\mu\text{g ml}^{-1}$, 126 $\mu\text{g ml}^{-1}$. Reactions were further incubated at 27 °C for 10, 40, 80 or 140 min, with antibody to Vam3 where indicated (**e**), and assayed for **d**, v-SNARE/t-SNARE complexes by co-immunoprecipitation, as described in Fig. 5a, or **e**, ALP activity after 140 min. Immunoblots were stained with antibodies against Nyv1, Vam7 or Vam3.

fusion reaction (Fig. 6c). As fusion is far slower than priming or docking in our reaction (Fig. 4a), this indicates that SNARE pairs might no longer be required at the fusion event. We therefore used a staged reaction to study the role of SNARE pairs at the fusion stage *per se*. Vacuoles from the Δt and Δv strains were incubated with cytosol for 40 min at 27°C to allow full SNARE pairing (Fig. 5), re-isolated, resuspended with cytosol, Sec17, ATP and increasing amounts of purified Sec18, and incubated for 10 min at 27°C. Aliquots were solubilized in detergent and analysed by co-immunoprecipitation with an antibody against Vam3, assayed for fusion that had occurred during these first incubations, allowed to complete the 180-min incubation and assayed for fusion, or mixed with

an antibody to the t-SNARE Vam3 for the remainder of the 180-min fusion assay. At the time of re-isolation, almost all the vacuoles had primed and docked and ~35% of the vacuoles had fused (Fig. 6e, black bars). Addition of Sec18 to the second incubation caused dissociation of the newly formed complex between Nyv1 and Vam3 (Fig. 6d, lanes 3–7). Even longer exposure of the Nyv1 immunoblot did not reveal any residual association with Vam3 (Fig. 6d, bottom panel, lanes 6, 7), showing that the dissociation was >98% as analysed by densitometry. Vam7 was also part of this complex and disassembled from Vam3 as did Nyv1. However, the fusion activity in the second incubation was not affected by the disassembly of the *trans*-SNARE complex regardless of the amount of Sec18 added (Fig. 6e; black bars, fusion at 50 min; grey bars, fusion at 180 min). Furthermore, addition of antibodies against Vam3 had no effect on fusion during this final incubation (Fig. 6e, white bars), although the SNAREs had been efficiently unpaired (Fig. 6d) and the SNARE complexes did not re-form later in the reaction (Fig. 6d, lanes 8–10). As SNARE pairs formed at the docking site can be targets of the ATP-dependent action of Sec17/Sec18, and as this unpairing does not affect the rate of fusion, SNARE pairs are not the proximal catalysts of the fusion event itself.

Conclusions

Vacuole docking occurs in two distinct steps (Fig. 7), a reversible Ypt7-dependent tethering and irreversible SNARE pairing. The *trans*-SNARE complex maintains specificity and directionality to the docking process but does not mediate the initial association of vacuoles. SNARE pairs formed in *trans* can be disassembled by Sec17/Sec18/ATP after docking without influencing the rate of fusion (Fig. 7), indicating that other factors may be required for the fusion reaction.

SNARE-independent tethering has also been described for transport from the endoplasmic reticulum to the Golgi¹¹, for the docking of Golgi vesicles at the end of mitosis²⁶ and for rabaptin-5- and Rab5-dependent tethering during endosome–endosome fusion²⁷. In agreement with these studies, we find that Ypt7 helps mediate vacuole–vacuole contact, as tethered vacuoles will disperse upon extraction of Ypt7 by Gdi1 (Figs 1b, 2, 3). However, Gdi1-mediated dispersal cannot occur after SNAREs pair, providing added stability to docking. The tethering of wild-type vacuoles requires priming²⁰. As Ypt7 is not a stable part of the vacuole SNARE complex²⁴ and tethering is a SNARE-independent reaction (Figs 2 and 3), a signal must be transmitted during or after priming that triggers the tethering reaction. This signalling is not required for fusion of vacuoles that lack an assembled SNARE complex, for which tethering can occur in a Sec17/Sec18/ATP-independent reaction (Fig. 3).

SNAREs have been implicated in both docking and fusion. We have now shown that SNAREs from opposing organelles actually form a *trans*-SNARE complex in a priming- and tethering-dependent reaction. Our data are consistent with studies of synaptic vesicles. Potent neurotoxins that cleave only unpaired SNAREs⁸ can efficiently block synaptic transmission^{28,29}. When neurotoxins are injected into squid nerve terminals, vesicles remain tethered or even accumulate further at the active zone¹³, and thus may be tethered without SNARE pairing. NSF and α -SNAP disassemble SNARE complexes on synaptic vesicles and the plasma membrane¹ and are needed to prime tethered exocytic vesicles³⁰; we suggest that this priming unpairs *cis*-SNARE complexes, activates the t-SNARE, and thereby leads to *trans*-SNARE pairing. In this model of neural transmission, tethering would precede NSF- and ATP-mediated priming, as already shown for endoplasmic-reticulum-derived vesicles at the Golgi¹¹. The concept of vesicles being tethered at the active zone without SNARE pairing is also supported by studies of *Drosophila melanogaster*, in which deletion of the genes encoding syntaxin (a t-SNARE) or synaptobrevin (a v-SNARE) still allows vesicles to tether to the active zone without the capacity for Ca²⁺-mediated fusion¹⁴.

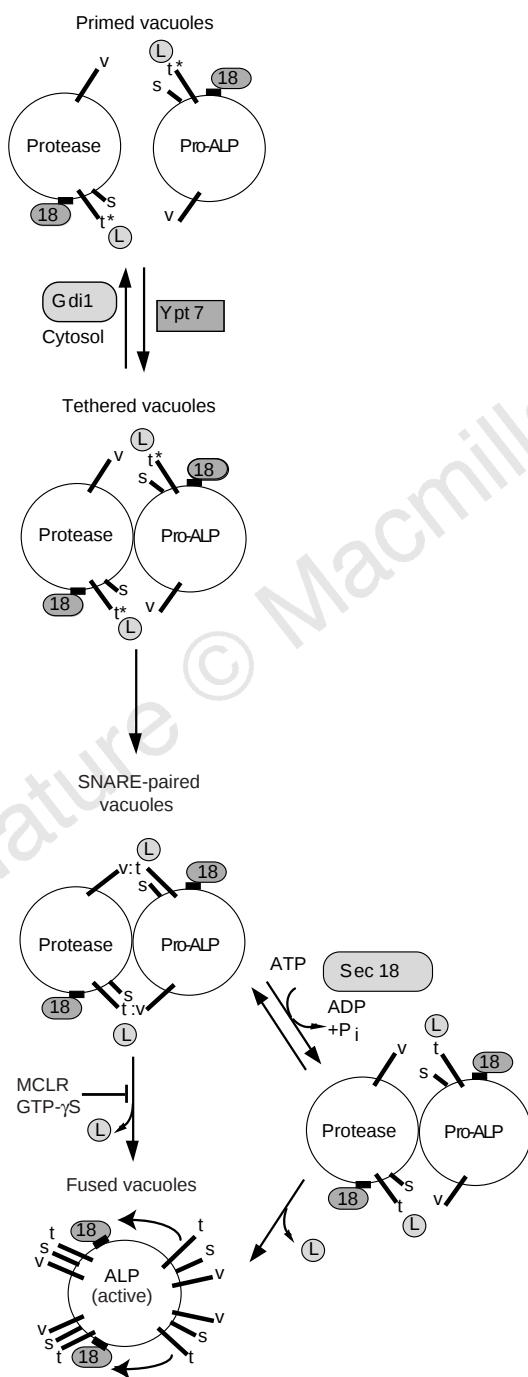


Figure 7 Working model for docking and fusion. Asterisks on the t-SNAREs indicate the activated state after priming. L, LMA1; s, SNAP-25 homologue (Vam7); t, t-SNARE; v, v-SNARE; 18, Sec18.

Purified SNAREs reconstituted into opposing liposomes promote lipid mixing, indicating that SNAREs may be the proximal catalysts of fusion¹⁵. However, our studies of homotypic vacuole fusion have revealed several inhibitors of the fusion step, such as MCLR, GTP- γ S and mastoparans, that are unlikely to act directly upon SNAREs. Furthermore, the dissociation of SNARE pairs in *trans* has no effect on the rate of fusion. Might sufficient SNARE pairs remain on Sec18-treated vacuoles to be the essential fusion catalyst? SNARE complexes are not abundant, being found in no more than 100 copies per vacuole (C.U., unpublished observations). Thus, a steady state of *trans*-SNARE pairs at the 2% level (Fig. 5) and an efficient disassembly (>98%) of these few SNARE pairs (Fig. 6) leaves far less than one SNARE pair per docked vacuole. Although it is possible that a labile subset of SNARE pairs might escape detection in detergent extracts, it is more likely that such lability would allow them to be dissociated by Sec18/Sec17/ATP.

What then is the role of SNARE pairs? They certainly function to stabilize the otherwise reversible association of tethered vacuoles. Although they might promote hemifusion, consistent with the fact that the fusion of docked vacuoles in which SNARE pairing has been disrupted remains resistant to Gdi1 or dilution (C.U., unpublished observations), SNARE pairs are no longer required for the MCLR-sensitive step that leads directly to complete membrane fusion and mixing of contents. We suggest that SNARE pairs signal downstream events that themselves catalyse fusion. These events may include docking-dependent calcium flux, Ca²⁺/calmodulin interactions and association with downstream effectors³¹. Other downstream events include release of LMA1, regulated by a phosphatase and opposing kinase¹⁹, and as yet unidentified targets of GTP- γ S and mastoparans. The availability of a detergent-mixed micellar extract of vacuole membranes, which can be reconstituted to form proteoliposomes that can perform authentic priming, docking and fusion²⁴, provides an avenue towards the enzymological resolution of the essential factors of these reactions. □

Methods

Reagents and yeast strains^{23,25,32,33}. SDS-polyacrylamide gel electrophoresis (SDS-PAGE), immunoblotting using chemiluminescence^{3,32}, purification of His₆-tagged Sec18 (ref. 33), co-immunoprecipitations³ and purification of IgGs¹⁸ were as described. Vacuoles³² were used immediately after isolation. Standard fusion reactions contained 3 μ g each of vacuoles from strains BJ3505 and DKY6281 with wild-type SNAREs, or 4 μ g of each 'mutant vacuole' (BJ3505 Δ nyv1 and DKY6281 Δ vam3 Δ) in reaction buffer with protease inhibitor cocktail³⁴. One unit of fusion activity is measured as 1 μ mol *p*-nitrophenol phosphate hydrolysed per min per microgram BJ3505 vacuole.

Proteoliposome fusion with intact vacuoles. Vacuoles from strain BJ3505 were detergent-solubilized and reconstituted into RPLs in the presence of pro-alkaline phosphatase (ALP) as described²⁴ (RPLs^{Pro}). Vacuoles from strain DKY6281 were incubated with RPLs^{Pro} (3 μ g each) in the presence of cytosol and ATP and the indicated inhibitors for 180 min at 27 °C (Fig. 2a). Mature phosphatase activity was measured.

Tethering assay using RPLs. RPLs were generated from BJ3505 vacuoles as above with mature ALP (10 μ g ml⁻¹; RPLs^M) but with 1 mg ml⁻¹ protein in the detergent extract and final sonication for 30–40 s (Fig. 2b). *In vitro* maturation of His₆-pro-ALP was by 100 μ g ml⁻¹ Proteinase A (Sigma) in 20 mM HEPES/KOH, pH 8.0, 0.5 M NaCl, 30 mM imidazole, pH 8.0, 1 mM dithiothreitol and 1.25% (w/v) *n*-octyl β -D-glucopyranoside for 10 min at 30 °C. The reaction was mixed with Ni-NTA agarose and matured alkaline phosphatase was purified as

described²⁴. For tethering, vacuoles (3 μ g) from strain BJ3505 were incubated with RPLs^M (10 μ l), 40 μ g ml⁻¹ Sec18, 1 mg ml⁻¹ cytosol, ATP and inhibitors where indicated for 45 min at 27 °C, then diluted with 570 μ l reaction buffer. Vacuoles with their tethered RPLs^M were isolated (10 min, 4,500g, 4 °C) and resuspended in 30 μ l PS buffer and ALP activity was measured.

Received 7 August; accepted 27 October 1998.

- Otto, H., Hanson, P. I. & Jahn, R. Assembly and disassembly of a ternary complex of synaptobrevin, syntaxin and SNAP-25 in the membrane of synaptic vesicles. *Proc. Natl Acad. Sci. USA* **94**, 6197–6201 (1997).
- Holthuis, J. C. M., Nichols, B. J., Druvakumar, S. & Pelham, H. R. B. Two syntaxin homologues in the TGN/endosomal system in yeast. *EMBO J.* **17**, 113–126 (1998).
- Ungermann, C., Nichols, B. J., Pelham, H. R. B. & Wickner, W. A vacuolar v-t-SNARE complex, the predominant form *in vivo* and on isolated organelles, is disassembled and activated for docking and fusion. *J. Cell Biol.* **140**, 61–69 (1998).
- Rothman, J. E. Mechanisms of intracellular membrane fusion. *Nature* **372**, 55–63 (1994).
- Ferro-Novick, S. & Jahn, R. Vesicle fusion from yeast to man. *Nature* **370**, 191–193 (1994).
- Hay, J. C. & Scheller, R. SNAREs and NSF in targeted membrane fusion. *Curr. Opin. Cell Biol.* **9**, 505–512 (1997).
- Hayashi, T. et al. Synaptic vesicle membrane fusion complex: action of clostridial neurotoxins on assembly. *EMBO J.* **13**, 5051–5061 (1994).
- Hayashi, T., Yamasaki, S., Nauenburg, S., Binz, T. & Niemann, H. Disassembly of the reconstituted synaptic vesicle membrane fusion complex *in vitro*. *EMBO J.* **14**, 2317–2325 (1995).
- Pfeffer, S. R. Transport vesicle docking: SNARE and associates. *Annu. Rev. Cell Dev. Biol.* **12**, 441–461 (1996).
- Novick, P. & Zerial, M. The diversity of Rab proteins in vesicle transport. *Curr. Opin. Cell Biol.* **9**, 496–504 (1997).
- Cao, X., Ballew, N. & Barlow, C. Initial docking of ER-derived vesicles requires Uso1p and Ypt1p but is independent of SNARE proteins. *EMBO J.* **17**, 2156–2165 (1998).
- Sapperstein, S. K., Lupashin, V. V., Schmitt, H. D. & Waters, M. G. Assembly of the ER to Golgi SNARE complex requires Uso1p. *J. Cell Biol.* **132**, 755–767 (1996).
- Hunt, J. M. et al. A post-docking role for synaptobrevin in synaptic vesicle docking. *Neuron* **12**, 1269–1279 (1994).
- Brodie, K. et al. Syntaxin and synaptobrevin function downstream of vesicle docking in *Drosophila*. *Neuron* **15**, 663–673 (1995).
- Weber, T. et al. SNAREpins: minimal machinery for membrane fusion. *Cell* **92**, 759–772 (1998).
- Weisman, L. S. & Wickner, W. T. Intervacuole exchange in the yeast zygote: a new pathway in organelle communication. *Science* **241**, 589–591 (1988).
- Conradt, G., Shaw, J., Vida, T., Emr, S. & Wickner, W. *In vitro* reactions of vacuole inheritance in *Saccharomyces cerevisiae*. *J. Cell Biol.* **119**, 1469–1479 (1992).
- Ungermann, C. & Wickner, W. Vam7p, a vacuolar SNAP-25 homolog, is required for SNARE complex disassembly, docking and fusion. *EMBO J.* **17**, 3269–3276 (1998).
- Xu, X., Sato, K. & Wickner, W. LMA1 binds to vacuoles at Sec18p (NSF), transfers upon ATP hydrolysis to a t-SNARE (Vam3p) complex, and is released during fusion. *Cell* **93**, 1125–1134 (1998).
- Mayer, A. & Wickner, W. Docking of yeast vacuoles is catalyzed by the ras-like GTPase Ypt7p after symmetric priming by Sec18p (NSF). *J. Cell Biol.* **136**, 307–317 (1997).
- Conradt, B., Haas, A. & Wickner, W. Determination of four biochemically distinct, sequential stages during vacuole inheritance *in vitro*. *J. Cell Biol.* **126**, 99–110 (1994).
- Haas, A., Conradt, B. & Wickner, W. G-protein ligands inhibit *in vitro* reactions of vacuole inheritance. *J. Cell Biol.* **126**, 87–97 (1994).
- Mayer, A., Wickner, W. & Haas, A. Sec18p (NSF)-driven release of Sec17p (α -SNAP) precedes docking and fusion of yeast vacuoles. *Cell* **85**, 83–94 (1996).
- Sato, K. & Wickner, W. Functional reconstitution of Ypt7 GTPase and vacuolar SNARE complex. *Science* **281**, 700–702 (1998).
- Nichols, B. J., Ungermann, C., Pelham, H. R. B., Wickner, W. & Haas, A. Homotypic vacuolar fusion mediated by v- and t-SNAREs. *Nature* **387**, 199–202 (1997).
- Lowe, M., Nakamura, N. & Warren, G. Golgi division and membrane traffic. *Trends Cell Biol.* **8**, 40–44 (1998).
- Stenmark, H., Vitale, G., Ullrich, O. & Zerial, M. Rabaptin-5 is a direct effector of the small GTPase Rab5 in endocytic membrane fusion. *Cell* **83**, 423–432 (1995).
- Blasi, J. et al. Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25. *Nature* **365**, 160–163 (1993).
- Jahn, R. & Niemann, H. Molecular mechanisms of clostridial neurotoxins. *Ann. NY Acad. Sci.* **733**, 245–255 (1994).
- Martin, T. F. J. Stages of regulated exocytosis. *Trends Cell Biol.* **7**, 271–276 (1997).
- Peters, C. & Mayer, A. Ca²⁺/calmodulin signals the completion of docking and triggers a late step of vacuole fusion. *Nature* **396**, 575–580 (1998).
- Haas, A. A quantitative assay to measure homotypic vacuole fusion *in vitro*. *Methods Cell Sci.* **17**, 283–294 (1995).
- Haas, A. & Wickner, W. Homotypic vacuolar fusion requires Sec17p (yeast α -SNAP) and Sec18p (yeast NSF). *EMBO J.* **15**, 3296–3305 (1996).
- Xu, Z. & Wickner, W. Thioredoxin is required for vacuole inheritance in *S. cerevisiae*. *J. Cell Biol.* **132**, 787–794 (1996).

Acknowledgements. This work was supported by a grant from the NIH and fellowships from the DFG (to C.U.) and the HFSP (to K.S.).

Correspondence and requests for materials should be addressed to C.U. (e-mail: Christian.Ungermann@Dartmouth.edu).

Galaxy disruption as the origin of intracluster light in the Coma cluster of galaxies

Michael D. Gregg* & Michael J. West†

* Physics Department, University of California, Davis, California 95616, USA, and Institute for Geophysics and Planetary Physics, Lawrence Livermore National Laboratory, L-413 Livermore, California 94550, USA

† Department of Astronomy and Physics, Saint Mary's University, Halifax, Nova Scotia B3H 3C3, Canada

Although the existence of a faint background of starlight in the core of the Coma cluster has been well established^{1,2}, its origin is uncertain. This vast sea of stars could have formed outside the galaxies, using gas left over from the time of the cluster's birth. Alternatively, it might be the accumulated debris generated by interactions between the galaxies over the lifetime of the cluster³. Here we report the discovery of three large, low-surface-brightness features in the Coma cluster, the most spectacular of which is a plume-like structure, 130 kpc long, in the cluster's heart. These structures will disperse over the next one to two billion years, thereby enhancing the general background light. If this epoch is typical, we argue that a significant fraction—perhaps even most—of the intracluster light results from a steady accumulation of tidal debris generated during galaxy–galaxy and galaxy–cluster interactions.

We have been using the Kitt Peak Burrell Schmidt 0.6-metre telescope to conduct a multi-colour survey of the Coma cluster and its environs over a $3^\circ \times 3^\circ$ area. The CCD (charge-coupled device) on the Schmidt telescope has $2''$ pixels and a field of view of 1.1° . In March 1996, we took a series of images centred midway between NGC4874 and NGC4889, the supergiant elliptical galaxies which dominate the centre of the Coma cluster.

Inspection of the final R-band image (Fig. 1) reveals the presence of three large, low-surface-brightness (LSB) features. Their locations and approximate sizes are indicated by the heavy solid lines labelled 1, 2 and 3 in Fig. 1. All of these objects can also be seen in the B, V and I frames, confirming their reality. All are much larger and more prominent than the recently reported slim arc of tidal debris in the Coma cluster⁴ (no. 4 in Fig. 1). Given the myriad images of the Coma cluster that have been obtained over the years, it is surprising that these features have escaped notice until now. There are at least two instances in the literature of images that show one or more of them clearly^{5,6}, but without any comment by the authors. To isolate the LSB features, we digitally 'cleaned' portions of the images to remove superposed galaxies and stars by interpolation with background replacement (Figs 2 and 3).

LSB object no. 1. The most striking of the features is a plume-like structure adjacent to NGC4874 (Figs 1 and 2). It extends for at least $4.5'$ and varies from $30''$ to $60''$ in width; it may well extend farther to the west, past the large elliptical galaxy NGC4864. Adopting a distance of 100 Mpc to Coma, the plume is ~ 130 kpc in length and varies from 15 to 30 kpc wide. The integrated apparent R-band magnitude is $R = 15.6 \pm 0.1$, though it may contain more material outside our adopted borders. The mean R-band surface brightness

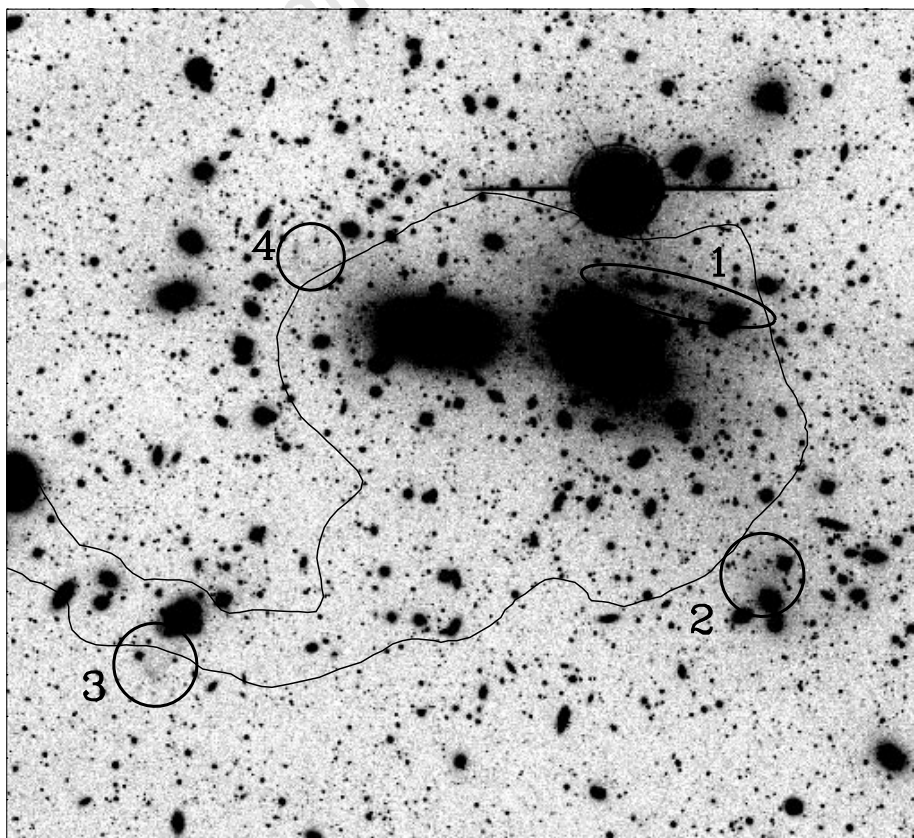


Figure 1 R-band image of the Coma cluster core showing the locations and approximate sizes of the low-surface-brightness features (heavy solid lines labelled 1–4) discussed here. The image is $37' \times 33'$; north is up, east to the left. Observing conditions were photometric with good seeing, and there was no Moon visible. The R-band integration totals 75 min; the B-, V-, and I-band images total 30–45 min. The images were reduced using standard procedures, and

combined using image masks to remove cosmetic defects on the CCD. Flatfielding using twilight images resulted in peak-to-peak sky flatness of $\sim 2\%$ in all colours. The thin solid line is one X-ray contour from Rosat observations¹¹ showing the extended ridge which includes LSB object no 3. The dark horizontal bar is charge bleeding on the CCD due to the bright star.

is $\mu_R = 25.7 \text{ mag arcsec}^{-2}$. The colours of the plume are consistent with those of normal elliptical galaxies of intermediate brightness, although the photometric uncertainties and removal of 'superposed' objects leave open the possibility of significant star formation in the LSB material.

LSB object no. 2. This feature is a pool of diffuse light adjacent to a group of galaxies including IC3957, IC3959 and IC3963 (left panel, Fig. 3). It is located near one end of a broader, extended swath of even lower surface brightness cluster background light⁵, linking it to NGC4874. This association makes it likely that the feature is within the Coma cluster core. The pool of material has a roughly circular distribution, $\sim 80''$ in diameter ($\sim 40 \text{ kpc}$). Its colours and surface

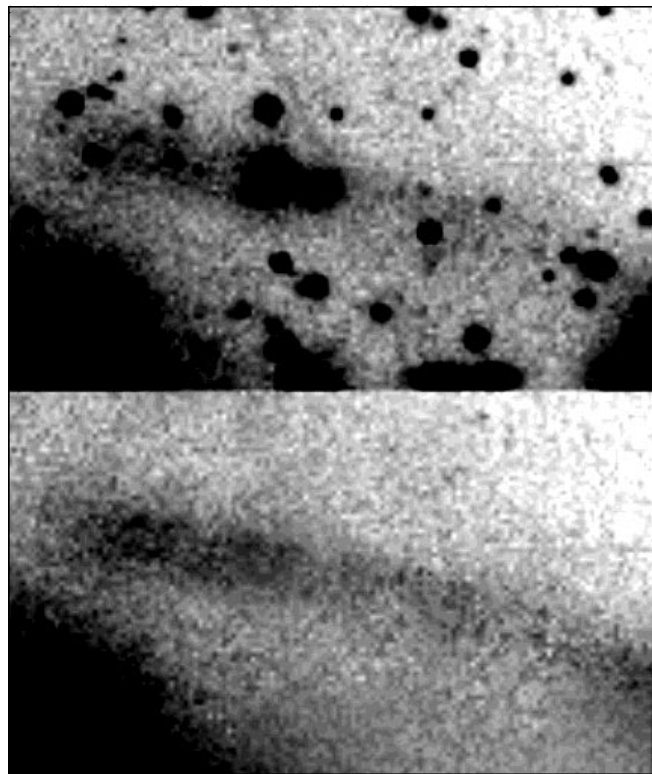


Figure 2 The plume feature before (top) and after (bottom) removing foreground objects. North is up, east is to the left. The sloping background from the giant elliptical NGC4874 was fitted and removed, and the remaining counts from the plume were summed to determine a total brightness. The plume is roughly $4.5' \times 1'$ in extent, with an integrated R-band magnitude $M_R = 15.6$ and mean R-band surface brightness of $\sim 26.3 \text{ mag arcsec}^{-2}$. Adopting a distance of 100 Mpc for the Coma cluster (equivalent to assuming $H_0 \approx 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$), the plume is roughly $130 \times 20 \text{ kpc}$ in extent, with an integrated absolute R-band magnitude of -19.4 . To estimate the colours, we restricted attention to an area of $80'' \times 40''$ where the surface brightness is highest, near the eastern end. The removal of embedded features was done in a consistent way in all colours. These inclusions may in fact be part of the tidally disrupted material, but it is impossible to determine this with certainty from the low-resolution Schmidt images; our conservative approach was to measure only the smooth underlying light belonging to the plume. Although the background can be determined to better than 1%, this contributes an uncertainty of 0.1–0.2 mag (10–20%) in the colours because of the faintness of the feature. The plume's colours are $B - V = 0.93$, $V - R = 0.57$ and $V - I = 1.17$, consistent with those of a normal elliptical galaxy. For comparison, the normal elliptical NGC3377 (ref. 25) has $B - V = 0.92$, $V - R = 0.58$, $V - I = 1.20$ and an absolute R-band magnitude²⁶ of -20.8 . These colours indicate little if any recent star formation associated with the tidal event. This constraint is weak, however, because the colours have been determined from a restricted area and the removed, apparent foreground objects may belong to the plume.

brightness are similar to those of the plume, with a total integrated magnitude 0.6 fainter.

LSB object no. 3. The third LSB feature is quite chaotic in appearance (right panel, Fig. 3). The lowest-surface-brightness portions of this object form a bridge to the outskirts of the Coma member galaxy NGC4911, where there is additional LSB material (inset, Fig. 3); the association places the feature in the Coma cluster. Several small knots are visible in the highest-surface-brightness portion, suggestive of star formation; but its colours are not very blue as would be expected in that case. Its overall photometric properties are similar to the other LSB features.

LSB object no. 4. This has already been described in detail by Trentham and Mobasher⁴ as a "giant low surface brightness arc of length $\sim 80 \text{ kpc}$ ". From our low-resolution Schmidt images, we cannot confidently classify this feature as a tidal arc rather than as a curious alignment of faint galaxies, nor can we analyse it in any detail. Its total length is $180''$ with a mean B-band surface brightness of $< 26.5 \text{ mag arcsec}^{-2}$ (ref. 4); in our R-band image, its average width is $\sim 5''$. Assuming typical galaxy colours, we estimate the integrated R-band magnitude as ~ 18 , roughly 5–10 times fainter than the other LSB features reported here, with a comparable mean surface brightness.

The LSB features all appear to be non-equilibrium configurations of stars deep within the gravitational potential well of the Coma cluster. We believe that they are best interpreted as transient features produced either by galaxy–galaxy interactions or by stripping of galaxies by the global cluster tidal field. Coma's large $\sim 1,000 \text{ km s}^{-1}$ velocity dispersion would seem to preclude strong galaxy–galaxy interactions as the energy transferred between systems is roughly proportional to the time spent in close proximity; typical high-speed encounters between galaxies will not produce significant disturbances⁷. Yet the presence of these LSB objects shows that Coma's galaxy population is undergoing vigorous dynamical evolution. There is now ample evidence from both optical and X-ray observations that the Coma cluster is a dynamically young system with several components in the throes of merging^{2,8–12}. The chaotic nature of subcluster mergers and the relatively low internal velocity dispersions of constituent subgroups, such as the recently identified "galaxy aggregates" in Coma¹³, will provide opportunities for lower-speed encounters, producing both galaxy–galaxy and galaxy–cluster tidal interactions.

Linear features such as the plume are routinely produced in numerical simulations of galaxy tidal interactions^{14,15}, although the plume's origin is hard to pinpoint from our data. It may be the result of interaction between the two elliptical-like nuclei embedded near its centre, though their large velocity difference ($> 1,100 \text{ km s}^{-1}$)¹⁶ makes it unlikely that they could have generated such a long tidal feature while remaining apparently close together¹⁴. Tides acting in the outer regions of NGC4864 (the large elliptical galaxy located near the plume's western end) could also produce the plume. Another possibility, perhaps less likely, is that the plume is the remains of an entire galaxy being disrupted by the global tidal field as it plunges into the cluster core, not unlike the dive into Jupiter of comet Shoemaker–Levy.

The amorphous structures of the LSB features 2–4 are not typical of what is seen in simulations, or examples on the sky, of tidal debris; nevertheless, tidal interaction is a natural explanation. LSB object no. 2 may be the product of strong tidal interaction within the subgroup containing IC3959 and IC3963; their velocity difference is only 240 km s^{-1} (ref. 16). NGC4911 and LSB object no. 3 lie within a $\sim 1\text{-Mpc}$ -long ridge of heightened X-ray emission (Fig. 1) which may have been produced by tidal disruption of a galaxy group on its recent passage through the cluster core¹¹. This suggests that the LSB object no. 3 may have originated during the same dynamical interaction that has brightened the X-ray emission. It could be the remains of a galaxy that has completely disrupted, or it may be material that has been stripped from the outskirts of NGC4911. The

object described by Trentham and Mobasher is less clearly associated with cluster member galaxies, but it too is best explained by tidal processes⁴.

With four examples of rather severe tidal interactions, one should expect many more less spectacular encounters to be taking place. The coarse-resolution Schmidt images, however, may be incapable of revealing them as such; LSB object no. 4 is detected, but its nature is not apparent. Our data suggest that there are many weaker tidal interactions and galaxy disturbances taking place in Coma, but finer images which reach at least as deep will be necessary to analyse them. If evidence for weaker interactions is found, it will help to account for the dearth of low-density dwarf galaxies in Coma's

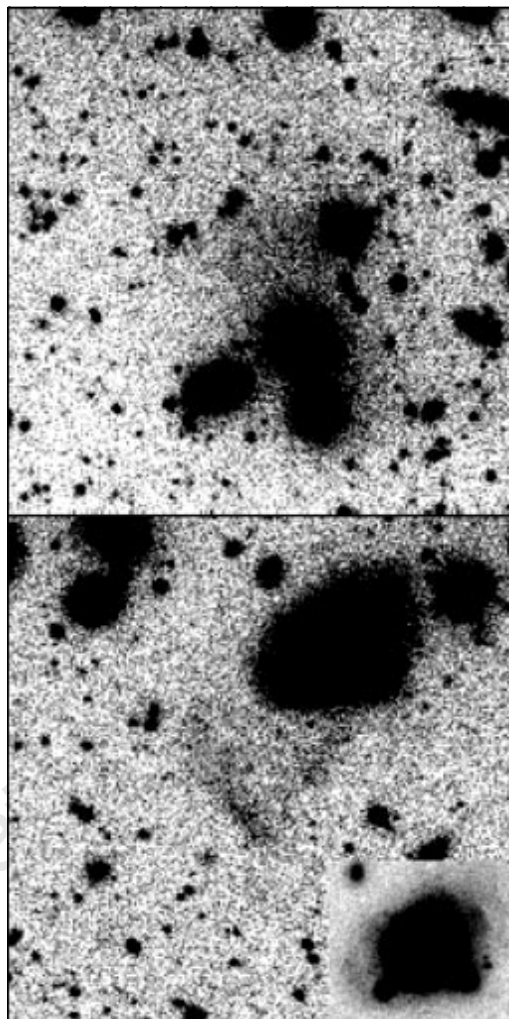


Figure 3 LSB objects no. 2 (left) and no. 3 (right) after removing foreground stars and galaxies. Each image is 400'' square; north is up, and east is to the left. The colours and absolute magnitude of object no. 2 are $B - V = 0.90$, $V - R = 0.43$, $V - I = 1.06$, $M_R = -18.8$; for no. 3 the values are $B - V = 0.93$, $V - R = 0.59$, $V - I = 0.89$, $M_R = -18.3$. Like the plume feature, these are similar to a normal elliptical galaxy and indicate little if any recent star formation, although here too the removed foreground objects may be part of the tidal material. The bright object immediately to the northwest of feature no. 2 is a Galactic star. The two objects immediately south of it are IC3959 and IC3963; their low velocity difference of 240 km s^{-1} may allow the strong tidal interaction which gives rise to the LSB material. The bright object to the northwest of feature no. 3 is NGC4911, a face-on spiral. The inset in the lower right is a lower-contrast display of NGC4911 showing the LSB material encircling its disk and linking it to LSB object no. 3. NGC4911 has a close S0 companion which is partly visible in the inset as a protrusion from the lower right of the disk. Their velocity difference is only 300 km s^{-1} , small enough to permit strong tidal interactions.

core¹⁷, as such galaxies are most susceptible to tidal disruption by the cluster's strong gravitational field, and also the anti-correlation between the luminosity of the brightest galaxy in a cluster and the size of the dwarf galaxy population¹⁸.

The LSB features will gradually disperse, either blending into the existing background of diffuse intracluster material or accreting onto the giant galaxies NGC4889 and NGC4874 at the bottom of Coma's gravitational potential well. All four tidal features are located within $\sim 0.5 \text{ Mpc}$ of the cluster centre and should dissipate in no more than a few cluster crossing times, about 1 to 2 Gyr. If this material contains quickly evolving massive stars, or even ongoing star formation, the use of brightest cluster galaxies as standard candles is further complicated^{19,20}. The very-highest-velocity material will contribute to the population of intergalactic stars and perhaps intergalactic globular clusters as well; evidence of such components may have been detected^{21–24} in other clusters. If we are viewing Coma at a typical moment in its history, then the presence of at least four large transient tidal debris piles indicates that a substantial amount of material could have been liberated from galaxies over the cluster's lifetime; every $\sim 1 \text{ Gyr}$, the equivalent of a galaxy of R-band absolute magnitude (M_R) ≈ -19 is being disrupted and accreted by the cluster core. Integrated over the 10–20-Gyr history of the Universe, the amount of stellar material added to the pool of intracluster light would be equivalent to an object with $M_R \approx -22$. This is about 20% of the total luminosity of either NGC4874 or NGC4889, and the interaction rate was almost certainly higher in the past. A significant portion of the mass of the central galaxies may consist of material accreted from tidally dismantled galaxies as the cluster evolves dynamically^{18,20}. The LSB objects now in Coma provide a vivid 'snapshot' of this process in action. $\square$

Received 23 April; accepted 16 September 1998.

1. Zwicky, F. *Morphological Astronomy* 48 (Springer, Berlin, 1957).
2. Biviano, A. in *A New Vision of an Old Cluster: Untangling Coma Berenices* (eds Mazure, A., Casoli, F., Durret, F. & Gerbal, D.) 1–8 (World Scientific, Singapore, 1998).
3. Merritt, D. Relaxation and tidal stripping in rich clusters of galaxies. II—evolution of the luminosity distribution. *Astrophys. J.* **276**, 26–37 (1984).
4. Trentham, N. & Mobasher, B. The discovery of a giant debris arc in the Coma cluster. *Mon. Not. R. Astron. Soc.* **293**, 53–59 (1998).
5. Welch, G. A. & Sastry, G. N. Photographic detection of "intergalactic" matter in the Coma cluster. *Astrophys. J.* **169**, L3–L5 (1971).
6. Secker, J. & Harris, W. E. Dwarf galaxies in the Coma cluster I. Detection, measurement and classification techniques. *Publ. Astron. Soc. Pacif.* **109**, 1364–1376 (1997).
7. Binney, J. & Tremaine, S. *Galactic Dynamics* (Princeton Univ. Press, 1987).
8. Fitchett, M. J. & Webster, R. L. Substructure in the Coma cluster. *Astrophys. J.* **317**, 653–667 (1987).
9. Mellier, Y., Mathez, G., Mazure, A., Chauvineau, B. & Proust, D. Subclustering and evolution of the Coma cluster. *Astron. Astrophys.* **199**, 67–72 (1988).
10. White, S. D. M., Briel, U. G. & Henry, J. P. X-ray archaeology in the Coma cluster. *Mon. Not. R. Astron. Soc.* **261**, L8–L12 (1993).
11. Vikhlinin, A., Forman, W. & Jones, C. Another collision for the Coma cluster. *Astrophys. J.* **474**, L7–L10 (1997).
12. Colless, M. & Dunn, A. M. Structures and dynamics of the Coma cluster. *Astrophys. J.* **458**, 435–454 (1996).
13. Conselice, C. J. & Gallagher, J. S. Galaxy aggregates in the Coma cluster. *Mon. Not. R. Astron. Soc.* **297**, L34–L38 (1998).
14. Dubinski, J., Mihos, J. C. & Hernquist, L. Using tidal tails to probe dark matter halos. *Astrophys. J.* **462**, 576–593 (1996).
15. Moore, B., Katz, N., Lake, G., Dressler, A. & Oemler, A. Jr Galaxy harassment and the evolution of clusters of galaxies. *Nature* **379**, 613–616 (1996).
16. NASA/IPAC Extragalactic Database available at (<http://nedwww.ipac.caltech.edu>) (1998).
17. Thompson, L. A. & Gregory, S. A. Dwarf galaxies in the Coma cluster. *Astron. J.* **106**, 2197–2212 (1993).
18. Lopez-Cruz, O., Yee, H. K. C., Brown, J. P., Jones, C. & Forman, W. Are luminous cD halos formed by the disruption of dwarf galaxies? *Astrophys. J.* **475**, L97–L101 (1997).
19. Tinsley, B. M. A first approximation to the effect of evolution on q_0 . *Astrophys. J.* **173**, L93–L97 (1972).
20. Ostriker, J. P. & Tremaine, S. D. Another evolutionary correction to the luminosity of giant galaxies. *Astrophys. J.* **202**, L113–L116 (1975).
21. Theuns, T. & Warren, S. J. Intergalactic stars in the Fornax cluster. *Mon. Not. R. Astron. Soc.* **284**, L11–L15 (1997).
22. Ferguson, H. C., Tanvir, N. R. & von Hippel, T. Detection of intergalactic red-giant-branch stars in the Virgo cluster. *Nature* **391**, 461–463 (1998).
23. Ciardullo, R., Jacoby, G. H., Feldmeier, J. J. & Bartlett, R. E. The planetary nebula luminosity function of M87 and the intracluster stars of Virgo. *Astrophys. J.* **492**, 62–73 (1998).
24. West, M., Cote, P., Jones, C., Forman, W. & Marzke, R. Intracluster globular clusters. *Astrophys. J.* **453**, L77–L80 (1995).
25. Gregg, M. D. Differential population synthesis of S0 galaxies. I—the data. *Astrophys. J. Suppl.* **69**, 217–232 (1989).
26. Sandage, A. & Tammann, G. A. *A Revised Shapley-Ames Catalog of Bright Galaxies* (Carnegie Institute of Washington, Washington DC, 1987).

Acknowledgements. Part of this work was done at the Institute of Geophysics and Planetary Physics, under the auspices of the US DOE by Lawrence Livermore National Laboratory. We acknowledge use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, Caltech, under contract with the NASA. Kitt Peak National Observatory, National Optical Astronomy Observatory, is operated by the Association of Universities for Research in Astronomy, Inc. (AURA), under cooperative agreement with the NSF. Observations were made with the Burrell Schmidt telescope of the Warner and Swasey Observatory, Case Western Reserve University. M.J.W. was supported by the NSERC of Canada.

Correspondence and requests for materials should be addressed to M.D.G. at LLNL (e-mail: gregg@igpp.llnl.gov).

Universality of rare fluctuations in turbulence and critical phenomena

S. T. Bramwell*, P. C. W. Holdsworth† & J.-F. Pinton‡

* Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK

† Laboratoire de Physique, Ecole Normale Supérieure, 46 Allée d'Italie, F-69364 Lyon cedex 07, France

A statistical treatment of three-dimensional turbulent flow continues to pose a challenge to theorists^{1,2}. One suggestion invokes an analogy with equilibrium phase transitions³. Here we approach this idea experimentally, presenting evidence of a strong analogy between the statistical behaviour of a confined turbulent flow and that of a model of the critical behaviour of a ferromagnet. Both systems experience large fluctuations limited only by the system size. We find that the power consumption measured in turbulent-flow experiments and the magnetization at the critical point of the ferromagnet have probability distributions of the same functional form, irrespective of Reynolds number on the one hand and system size on the other. The distributions both have non-gaussian tails that characterize the large-amplitude fluctuations. In this region, the scaled distributions for the two systems collapse onto a single universal curve over at least four orders of magnitude. This suggests a basic similarity in the finite-size corrections to the fluctuation statistics in the limit of infinite system size (for the magnetic system) or infinite Reynolds number (for turbulent flow).

The power consumption of turbulent flow is an important practical quantity in fluid mechanics, which is related, for example, to the drag experienced by a moving body. In particular, large amplitude fluctuations in this quantity, although rare, can have serious practical consequences. We have previously described an experiment⁴ that measures the temporal fluctuations in the power consumption of the turbulent flow produced in an enclosed air gap between two counter-rotating disks⁵. The two important constraints of this system are that the flow is spatially confined in a cylindrical vessel, and that the integral Reynolds number is fixed at a constant value. The Reynolds number Re characterizes the complexity of the fluid motion; it is defined as the ratio of the nonlinear to dissipative forces. For this experiment, it takes the value $Re = L^2\Omega/\nu$, where L and Ω are the radius and rotation rate of the disks, respectively, and ν is the fluid's kinematic viscosity. Re is fixed by controlling the electric motors that drive the disks, such that their rotation rate Ω is constant. This means that the rate of energy consumption of the flow $P(t)$, measured as the power delivered by the driving motors, fluctuates in time. The probability density function $Q_P(P)$ for the fluctuations in $P(t)$ has been measured⁴ for different Reynolds numbers. The mean value, $\bar{P}$, and the r.m.s. half-width, σ_P , of the distribution depend on Re , but the functional form of the distribution does not. A general method of confirming that different normalized probability distributions $f_n(x)$ (where n is 1, 2, 3 ...) have the same functional form is to plot $\sigma_n f_n(x)$ versus $(x - \bar{x}_n)/\sigma_n$, which brings the curves for different n into coincidence. The test is

applied to the turbulence data in Fig. 1a, where $\sigma_P Q_P(P)$ is plotted against $(P - \bar{P})/\sigma_P$ for different Re . Two important results can be deduced from Fig. 1a: first, the sets of measurements do indeed obey the same statistics, and second, the statistics are non-gaussian. In the figure, the rescaled curves are plotted on a logarithmic scale to emphasize the tail of the distribution.

The magnetic system under consideration is the spin-wave approximation to the two-dimensional XY model, which has proved to be a convenient system for calculating critical behaviour in the finite-size regime⁶. It is defined by the hamiltonian $H = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j)$, where J is the ferromagnetic coupling constant and θ_i the angle of orientation of the classical spin vector S_i , constrained to lie in a plane. The summation is over nearest neighbours and the spins are set on a square lattice with periodic boundary conditions. The spin-wave approximation, which is valid at low temperature, is obtained by expanding the cosine interaction to the quadratic term. Magnetic ordering is described by the instantaneous scalar order parameter:

$$M = \frac{1}{N} \sqrt{\left(\sum_{i=1,N} S_i \right)^2}$$

which has a probability distribution $Q_M(M)$, with mean value $\langle M \rangle$ and standard deviation σ_M . In Fig. 1b, we show the Monte Carlo results of ref. 6. As in the turbulence case, we observe universal statistics, this time as a function of N and T .

The central empirical result of this work is presented in Fig. 1c, where the probability density functions of the turbulence experiment and the critical system are overlaid on the same graph. These two apparently disconnected quantities obey essentially the same, universal, non-gaussian statistics. The high end of the distribution has a gaussian shape. Near the centre, small systematic discrepancies are found, both within and between the two sets of curves. In the magnetic case, these are $O(1/N)$ corrections to universality⁷, and in the turbulence case, they vary systematically with Re . The main region of interest, however, is the distinctive exponential tail towards the low end of the distribution, where the data coincide over four orders of magnitude, for a span of eight standard deviations. This result emphasizes the fact that in practical applications it is unsafe to assume that such fluctuations of a turbulent system are gaussian. For example, it can be deduced from the curve that the probability of a rare fluctuation of greater than six standard deviations from the mean is still 2×10^{-4} , as opposed to 1×10^{-9} in the gaussian case.

To see how the overlap of the two distributions can occur, we first observe that any possible analogue between the two systems must occur at a critical point; a singular point in the thermodynamic phase diagram, which is characterized by fluctuations that are correlated over a divergent length scale⁸. In a thermodynamic system, the driving force for the fluctuations is on the atomic scale. All trace of fluctuations on larger scales disappears everywhere except at the critical point, where they are maintained and grow right up to the macroscopic scale. In the turbulent fluid, the driving force is on the macroscopic scale, and the fluctuations are driven downward in scale to a dissipation length η , below which the system is uniform. Therefore, as in the critical system, each one of these length scales is important for establishing both the mean value of the dissipated power and the fluctuations about the mean.

Away from a critical point, fluctuations in a global quantity such as M can generally be treated as a small perturbation about the extremely precise mean value that defines equilibrium⁹. The standard deviation of the distribution, σ_M , is then proportional to $1/\sqrt{N}$. (This is a direct result of the statistical independence of the internal degrees of freedom and is a cornerstone of statistical thermodynamics. As N becomes large, the distribution becomes vanishingly narrow and is represented by a gaussian function, to an excellent approximation. In contrast, at a critical point the system is

scale-free or 'self-similar'¹⁰, meaning that on moving from one length scale to another, the form and intensity of the fluctuations remains unchanged and the scale invariance is only interrupted on arriving at the characteristic size L of the system¹¹. The fluctuations over different length scales all contribute to the fluctuations in M , with the result that $Q_M(M)$ is much broader; σ_M no longer varies as $1/\sqrt{N}$ and the distribution becomes non-gaussian.

The two-dimensional XY model provides an excellent example for studying critical behaviour as it has a continuous line of critical points over a finite range of temperature^{12,13}. At any one of these points, a renormalization-group treatment converges onto the spin-wave approximation¹⁴, which therefore describes the critical behaviour perfectly at all temperatures. The scale independence can be illustrated within the spin wave approximation, where one can easily calculate $\langle M \rangle$ and σ_M (refs 6, 15): $\langle M \rangle = (1/2N)^{k_B T/8\pi J}$, $\sigma_M = aT\langle M \rangle$, $a \approx 0.04$. In the thermodynamic limit, $N \rightarrow \infty$, the order parameter $\langle M \rangle$ is strictly zero¹⁶. However, the corrections to the thermodynamic limit disappear more slowly than the $\sim 1/\sqrt{N}$

variation discussed above^{11,17,18} and the ratio $\sigma_M/\langle M \rangle$ is independent of system size. Further, the self-similar nature of the fluctuations around this mean value is expected to reflect itself in the kind of scale independence for Q_M observed numerically in Fig. 1b (refs 19, 20). We have recently proved this explicitly⁷, showing that $Q_M(M)$ is indeed a universal function, independent of both N and T .

We now argue that there is a connection between system size and Reynolds number that leads to an analogy, at a stochastic level, between the magnetic model and the turbulent fluid. The fluid motion is described by the three-dimensional Navier–Stokes equation and turbulence has its origin in the nonlinear term, whose importance compared to the smoothing viscous forces is given by the Reynolds number. In our experiment, motion is created at large scales where energy is fed into the flow and transferred by the nonlinearity to smaller scales where the viscous forces become progressively more important. At a crossover length scale η , they dominate; energy dissipation exceeds energy transfer and the cascade terminates. Turbulence can thus be viewed as a super-

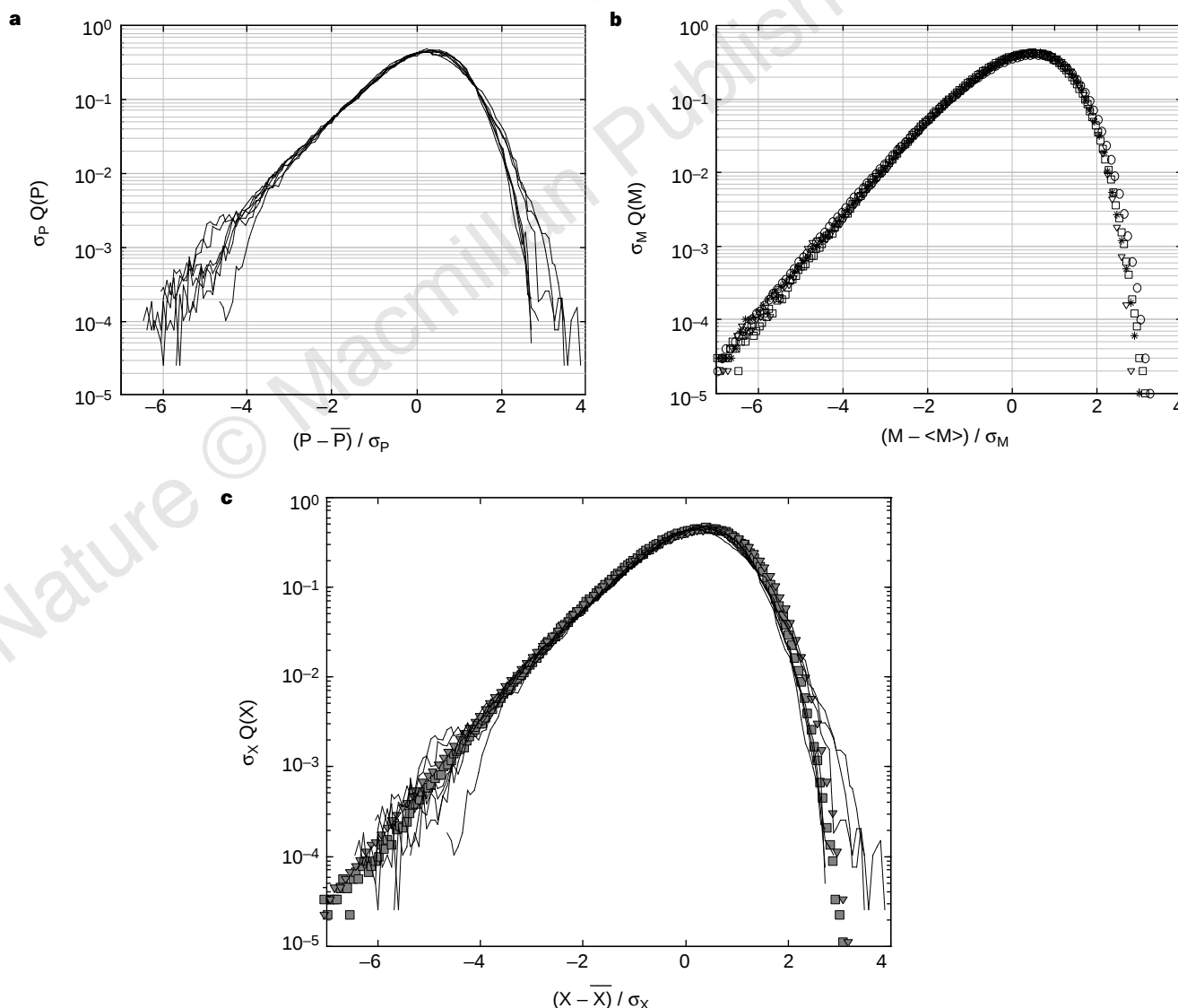


Figure 1 Universality of global fluctuations. Probability density function of power fluctuations, $\sigma_P Q_P$ in the turbulent closed flow produced in the gap between two counter-rotating disks (from ref. 4). Data are shown for rotation rates Ω of 25, 30, 35, 40 and 45 Hz. Using the reduced variable $(P - \bar{P})/\sigma_P$, the data measured at different Reynolds number, $Re \propto \Omega$, collapse onto a single curve. **b**, Probability density for the instantaneous magnetization from Monte Carlo simulation of the

spin wave approximation of the 2D-XY model at $k_B T/J = 0.5$, for $N = 100$ spins (circles), $N = 1,024$ spins (stars), $N = 10,000$ spins (triangles) and $k_B T/J = 1.0$ for $N = 1,024$ spins (squares)⁶. The data collapse onto a single universal curve for the reduced variable $(M - \langle M \rangle)/\sigma_M$. **c**, Data from **a** and **b** superimposed; where X is the power P or magnetization M , and $\bar{X} = \bar{P}$ or $\langle M \rangle$.

position of moving eddies on a wide range of scales. The smallest one is the dissipation length η . In a closed flow, the largest one is equal to the system size L . We emphasize the difference with open flows where eddies escape through the open boundaries and the power consumption has gaussian fluctuations⁴. The effective number of degrees of freedom contributing to the turbulence motion can be estimated as $N = (L/\eta)^3$ (ref. 21). Using Kolmogorov mean-field arguments, N can be related to the Reynolds number through $L/\eta = \text{Re}^{4/3}$ (ref. 1). Hence any system of finite Reynolds number can be considered as containing a finite number of degrees of freedom. Given that many scales are important, the simplest scenario is a self-similar one. Indeed, both Kolmogorov's original hypothesis about the energy cascade and subsequent corrections to include intermittency effects assume that the energy transfer per unit mass on scale l , ω_l , has a scale-invariant behaviour (for example, $\langle \epsilon_l^p \rangle \propto l^{p(p)}$), as confirmed by experiment^{1,22}. In a closed turbulent flow, the scale invariance holds up to the system size. The experimental evidence presented here suggests that this is sufficient to provide a complete analogy with the critical magnetic system and the same kind of universality for $Q_p(P)$ as for $Q_M(M)$.

A consequence of this interpretation is that the fluctuations in the dissipated power can be thought of as a finite-size correction to the result at infinite Reynolds number. The $\text{Re} = \infty$ case is by definition inaccessible for any experiment studying turbulence in a closed flow, as there is a well defined upper and lower length scale between which fluctuations are important. For very small values of Re , one might expect to observe transient effects that depend on the details and dimensions of the experiment, but once some threshold has been exceeded, $Q_p(P)$ should be independent of Re , however large it may be.

As many length scales are important at a critical point, the microscopic details often get washed away and the critical behaviour of apparently radically differing systems can be the same. Only large scale details, such as the symmetry of the hamiltonian and the dimension, are important, and systems with the same critical behaviour are grouped together in 'universality classes'⁸. The probability distribution function for all systems falling in the same universality class should therefore have the same universal form. However, there is no *a priori* reason to expect that the same should be true in going from one universality class to another. It is therefore surprising to find that these two functions are so similar. We do not believe that we have miraculously fallen onto the correct universality class for the turbulent fluid experiment. Rather, it is likely that, for certain universality classes, the departure from gaussian behaviour at a critical point is described to an excellent approximation by the spin-wave limit of the two-dimensional XY model, with the fine details that characterize and separate the universality classes being concentrated in the central part of the distribution function, or otherwise being hidden by experimental errors.

Our results indicate that the universality observed in the turbulence experiment can be explained in terms of a self-similar structure of fluctuations, just as in a finite critical system. This analogy provides an important new experimental application of finite size scaling approaches to a critical point²⁰ and it suggests a new range of experiments to characterize turbulent flow. Finally, it provides a systematic method of predicting the probability of rare fluctuations in a confined turbulent system. $\square$

Received 26 March; accepted 10 September 1998.

1. Frisch, U. *Turbulence* (Cambridge University Press, Cambridge, 1995).
2. Monin, A. S. & Yaglom, A. M. *Statistical Fluid Mechanics* (MIT Press, Massachusetts, 1975).
3. Eyink, G. L. & Goldenfeld, N. Analogies between scaling in turbulence, field theory and critical phenomena. *Phys. Rev. E* **50**, 4679–4683 (1994).
4. Labbé, R., Pinton, J.-F. & Fauve, S. Power fluctuations in turbulent swirling flows. *J. Phys. II (France)* **6**, 1099–1110 (1996).
5. Zandberg, P. J. & Dijkstra, D. Von Kármán flows. *Annu. Rev. Fluid Mech.* **19**, 465–491 (1987).
6. Archambault, P., Bramwell, S. T. & Holdsworth, P. C. W. Magnetic fluctuations in a finite two-dimensional XY model. *J. Phys. A: Math. Gen.* **30**, 8363–8378 (1997).
7. Archambault, P. *et al.* Universal magnetic fluctuations in the 2D-XY model. *J. Appl. Phys.* **83**, 7234–7236 (1998).

8. Wilson, K. G. & Kogut, J. The renormalization group and the ϵ expansion. *J. Phys. Rep. C* **12**, 75–200 (1974).
9. Landau, L. D. & Lifshitz, E. M. *Statistical Physics* Vol. 1 p.9 (Pergamon, Oxford, 1980).
10. Binney, J. J., Dowrick, N. J., Fisher, A. J. & Newman, M. E. J. in *The Theory of Critical Phenomena* 20 (Clarendon, Oxford, 1992).
11. Goldenfeld, N. Lectures on phase transitions and the renormalization group. *Frontiers Phys.* **85**, 279–282 (1992).
12. Kosterlitz, J. M. & Thouless, D. J. Metastability and phase transitions in two-dimensional systems. *J. Phys. C* **6**, 1181–1203 (1973).
13. Berezinskii, V. L. Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group I. Classical systems. *Sov. Phys. JETP* **32**, 493–500 (1971).
14. José, J. V., Kadanoff, L. P., Kirkpatrick, S. & Nelson, D. R. Renormalization, vortices, and symmetry-breaking perturbations in the two-dimensional planar model. *Phys. Rev. B* **16**, 1217–1241 (1977).
15. Tobochnik, J. & Chester, G. V. Monte Carlo study of the planar spin model. *Phys. Rev. B* **20**, 3761–3769 (1979).
16. Mermin, N. D. & Wagner, H. Absence of ferromagnetism or antiferromagnetism in one or two-dimensional isotropic Heisenberg models. *Phys. Rev. Lett.* **17**, 1133–1136 (1966).
17. Mermin, N. D. Absence of ordering in certain classical systems. *J. Math. Phys.* **8**, 1061–1064 (1967).
18. Bramwell, S. T. & Holdsworth, P. C. W. Magnetization and universal subcritical behaviour in two-dimensional XY magnets. *J. Phys. Condens. Matter* **5**, L53–59 (1993).
19. Binder, K. in *Computational Methods in Field Theory* (eds Gauslever, H. & Lang, C. B.) 59 (Springer, Berlin, 1992).
20. *Finite-Size Scaling, Current Physics Sources and Comments* Vol. 2 (ed. Cardy, J. L.) (North-Holland, 1990).
21. Landau, L. D. & Lifshitz, E. M. *Fluid Mechanics* 123 (Pergamon, Oxford, 1959).
22. Castaing, B., Gagne, Y. & Hopfinger, E. J. Velocity probability density functions of high Reynolds number turbulence. *Physica D* **46**, 177–200 (1990).

Acknowledgements. We thank P. Archambault, S. Fauve, J.-Y. Fortin, R. Labbé and S. Peysson for collaborations leading to refs 4, 6 and 7, and C. Baudet, J. L. Cardy, B. Castaing, M. J. Harris, J. Kurchan and Z. Rač for discussions.

Correspondence and requests for materials should be addressed to P.C.W.H. (e-mail: pcwh@enslapp.ens-lyon.fr).

Observation of 'third sound' in superfluid ³He

A. M. R. Schechter, R. W. Simmonds, R. E. Packard & J. C. Davis

Department of Physics, University of California, Berkeley, California 94720, USA

Waves on the surface of a fluid provide a powerful tool for studying the fluid itself and the surrounding physical environment. For example, the wave speed is determined by the force per unit mass at the surface, and by the depth of the fluid¹: the decreasing speed of ocean waves as they approach the shore reveals the changing depth of the sea and the strength of gravity. Other examples include propagating waves in neutron-star oceans² and on the surface of levitating liquid drops³. Although gravity is a common restoring force, others exist, including the electrostatic force which causes a thin liquid film to adhere to a solid. Usually surface waves cannot occur on such thin films because viscosity inhibits their motion. However, in the special case of thin films of superfluid ⁴He, surface waves do exist and are called 'third sound'. Here we report the detection of similar surface waves in thin films of superfluid ³He. We describe studies of the speed of these waves, the properties of the surface force, and the film's superfluid density.

Superfluid ³He can be described by a 'two fluid' model⁴ where the liquid is considered to be two interpenetrating fluids: a normal fluid component and a superfluid component. Each component has a mass density fraction, ρ_n/ρ and ρ_s/ρ respectively, and is governed by a different equation of motion. In particular, the superfluid component flows without viscosity while the normal component experiences viscous drag. This 'two fluid' nature allows several different types of acoustic phenomena to exist. For example, oscillations with both components moving in-phase are called first sound, while out-of-phase oscillations are called second sound.

Third sound^{5,6} is the name given to a surface wave travelling on a thin superfluid film. Here, the superfluid component oscillates parallel to the substrate while the normal-fluid component is held stationary by viscosity, as shown schematically in Fig. 1a. A typical wave amplitude is 0.1 nm. In the simplest case, the speed of

a third-sound wave takes the form⁵;

$$c_3 = \sqrt{\frac{\langle \rho_s \rangle}{\rho}} F \delta \quad (1)$$

where δ is the film thickness, $\langle \rho_s \rangle$ is the superfluid density averaged over δ , ρ is the density of the liquid, and F is the van der Waals force per unit mass at the film's surface. A possible additional term in equation (1) associated with specific entropy and temperature⁷ is negligible for our ^3He films.

Third sound has proven to be a powerful probe of fundamental phenomena in superfluid ^4He . Examples include measurements of the physics of two-dimensional phase transitions⁸, the layered nature of ^3He - ^4He films⁹, the non-wetting of alkali metals¹⁰, the generation of vortices and high circulation states¹¹, and many other phenomena¹²⁻¹⁶. The number and significance of these applications in ^4He has stimulated strong interest in whether similar surface waves can exist in films of superfluid ^3He ^{17,18}.

Several studies have explored the conditions for superfluidity in films of ^3He (refs 19-24), and have shown that it is suppressed when the film thickness is near the characteristic healing length, ξ . This is caused by breaking of p -wave Cooper pairs during scattering from the microscopically rough substrates. Because for ^3He $\xi = 65$ nm (at temperature $T = 0$ and zero pressure), we use films with thicknesses near 100 nm.

To search for third sound in ^3He we use the apparatus shown schematically in Fig. 1b. The superfluid ^3He film resides on the flat top surface of a polished copper disk. The basic experimental idea is to apply oscillating electrical forces to this film thereby exciting standing waves of third sound, and then to detect capacitively the surface motion due to these standing waves.

We performed test experiments in this apparatus¹⁸, using superfluid ^4He with films of thickness near 100 nm, which indicated that the perimeter of the substrate reflects the surface waves. Further, the spectrum of ^4He third sound is consistent with a 'fixed' boundary condition; that is, the surface displacement is zero near the substrate perimeter. This agrees with several similar experiments, including the first study⁶ of third sound in ^4He .

For a film of superfluid ^3He , solution of the wave equation with fixed boundary conditions at the substrate perimeter indicates that resonant standing-wave modes can exist. These modes are analogous to those in the membrane of a drum head, and their frequencies f are related to the wave speed c_3 by

$$f_{m,n} = c_3 \frac{\alpha_{m,n}}{2\pi R} \quad (2)$$

where the integers m and n are called the mode numbers, R is the radius of the disk, and $\alpha_{m,n}$ is the value of the n th zero of the Bessel function $J_m(x)$; $\alpha_{0,1} = 2.405$, $\alpha_{1,1} = 3.832$, $\alpha_{0,2} = 5.520$, $\alpha_{1,2} = 7.016$, $\alpha_{0,3} = 8.654$ specify the five lowest frequency modes to which this experiment is sensitive. Tests with superfluid ^4He in the present cell have identified resonances for these five modes.

In a typical series of superfluid ^3He measurements, the cell is cooled below the temperature where the film becomes superfluid, T_c^{film} , with a given film of thickness δ on the substrate. The driving-force frequency is slowly swept from 0.2 Hz to 8 Hz, and the amplitude of the surface response at each frequency is measured at the detector. We repeat this frequency sweep at a range of temperatures. Figure 2a shows the result of a series of such measurements for a film of $\delta = 233$ nm, between 0.32 mK and 0.79 mK. Several resonant modes, whose frequencies vary with changing temperature, are apparent. Similar spectra were acquired at film thicknesses, δ , of 92, 122, 170, 174, 189, 221, 252, 264 and 281 nm. From these, it is clear that well-defined surface-wave modes do exist in superfluid ^3He , and that we can identify their resonant frequencies and quality factors. In addition, the frequencies of these modes fall with rising temperature, and the phenomena disappear

for $T > T_c^{\text{film}}$. These data provide, to our knowledge, the first evidence for the existence of third sound in superfluid ^3He .

We can use these third-sound signals to probe the physics of the superfluid ^3He films by measuring the wave speed c_3 , and by using c_3 to obtain the average superfluid density $\langle \rho_s \rangle / \rho$. First, in order to assign mode numbers to the peaks shown in Fig. 2a, their frequencies are plotted versus temperature (Fig. 2b). We interpret the temperature dependence of the frequencies in terms of the simplest available model, using the following assumptions. First, equations

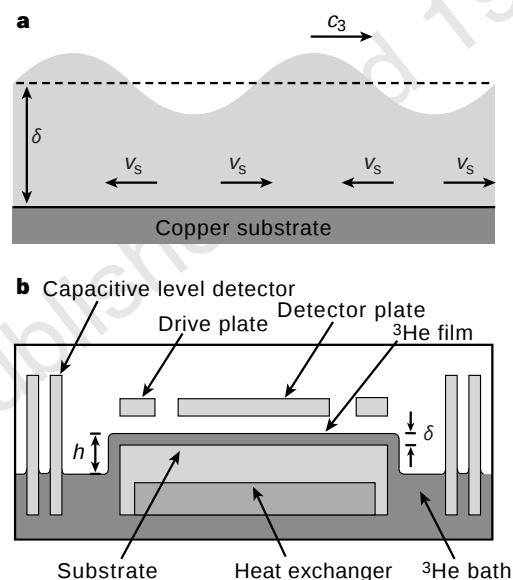


Figure 1 Schematic representations of a third-sound wave and of the experimental apparatus. **a**, The profile of a third-sound wave in which the 'normal-fluid' fraction of the film is kept in place by viscosity and the 'superfluid' fraction oscillates parallel to the substrate, with velocity v_s . This causes local variations in the film thickness while allowing a surface wave to propagate at speed c_3 . **b**, A thick horizontal copper disk (radius $R = 19.1$ mm) is positioned in a container of superfluid ^3He , its polished top surface forming the substrate for the superfluid film. From a fill line at the bottom of the cell, liquid ^3He is introduced to a distance h below the top of the disk. A film of thickness δ forms on the substrate when the temperature falls below the critical temperature for film superfluidity, T_c^{film} (which is less than the bulk superfluid critical temperature $T_c = 0.929$ mK). Two plates are placed above the substrate: a 10.1-mm-radius disk and an annular ring (13.6-mm inner radius and 17.8-mm outer radius). A grounded electrostatic shield (not shown) is positioned between the two electrodes. The geometry of the electrode assembly is designed to minimize the creation of standing waves of third sound on its surfaces. The plates are ~ 32 μm above the substrate, and are believed to be tilted so that one side is about 20 μm above the other side. This provides sensitivity to azimuthally asymmetric modes on the film. The outer annular electrode is connected to an oscillating voltage of typical amplitude 1 V superimposed on a larger d.c. voltage, typically 10 V. Due to the dielectric constant of the liquid ($\epsilon = 1.0426$) this produces an a.c. driving force on the film's surface. The central disk-shaped electrode is connected to a capacitance bridge that monitors film thickness changes which are occurring synchronously with the a.c. excitation. The system's resolution for average film surface displacement is ~ 30 pm Hz^{-1/2}. A coaxial cylindrical capacitor, mounted vertically, measures the height h from the free surface of the bulk liquid to the substrate. The apparatus is designed for the generation and detection of third sound in films with a range of film thicknesses near 100 nm, in a cell suitable for refrigeration down to $T < 300$ μK . The ^3He bath is cooled by a heat exchanger connected to a nuclear demagnetization refrigerator, and its temperature is measured using pulsed NMR on ^{195}Pt . As the substrate itself is isolated from electrical ground, thermal contact from it to the bath is made via a second heat exchanger, installed in the bottom of the substrate before polishing. The substrate was machined with a regular cutting tool and then polished to an apparent mirror finish by lapping. A series of diminishing grit sizes down to 0.25 μm were used.

(1) and (2) are combined to show that the mode frequencies are proportional to $(\rho_s/\rho)^{1/2}$. Second, using a Landau–Ginzburg model²⁴ which states that $\langle\rho_s\rangle/\rho \propto (1 - T/T_c^{\text{film}})$, we find that the mode frequencies $f_{m,n} \propto (1 - T/T_c^{\text{film}})^{1/2}$. We define T_c^{film} to be the measured temperature at which the film's motion in response to applied voltages disappears. The proportionality constant is identified by fitting the predicted frequencies for the third radial mode ($m = 0, n = 3$) to the frequencies of the (0,3) mode in the data. The predicted frequencies for the lower-frequency modes are then determined, and are shown as the solid lines in Fig. 2b.

The observed spectra clearly have features in common with the predictions of this simple model, allowing us to identify the mode numbers for several peaks. But other peaks are not in as good agreement, indicating that the surface waves are revealing new complexities in the physics of the ^3He film. For example, as T decreases we observe a behaviour involving the mixing and splitting of several lower-frequency modes, which is so far unexplained. Similar behaviour, both in the general agreement of several modes with the model, and in the unexpected splittings, have been observed for all other measured film thicknesses.

We use the observed frequencies of the (0,3) mode—because this mode is clearly identifiable for all films and temperatures—to determine c_3 . We believe that the (0,3) is most reliable because geometrically it has the strongest overlap with the drive capacitor

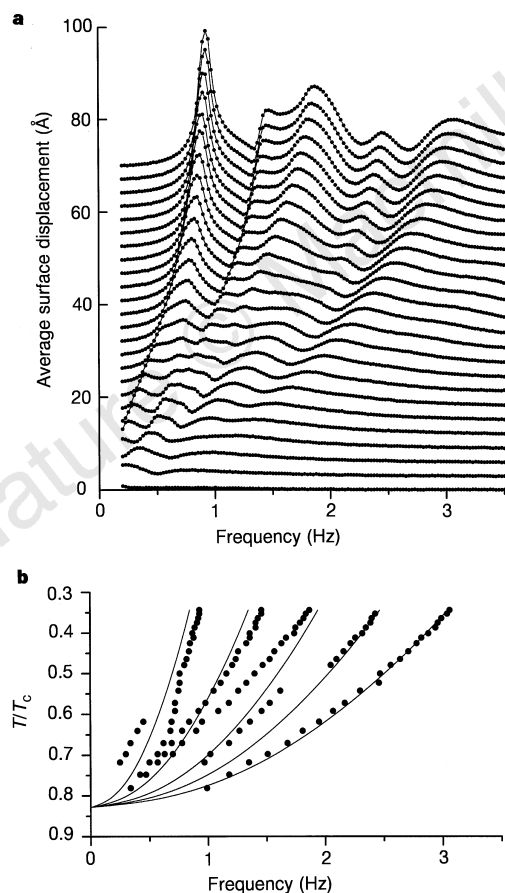


Figure 2 ^3He third-sound spectra, and the resonant frequencies as a function of temperature. **a**, A typical series of third-sound spectra for several different temperatures with $\delta = 233$ nm. The temperature ranges from 0.79 mK, just below the onset of superfluidity at T_c^{film} for the lowest curve, to 0.32 mK for the highest curve, with approximately even temperature steps. The curves are shifted vertically from each other for clarity of presentation. Similar spectra were acquired at film thicknesses $\delta = 92, 122, 170, 174, 189, 221, 252, 264$ and 281 nm. **b**, Frequencies of the observed peaks versus T/T_c . The solid lines represent the expected distribution of the peaks from the simple model described in the text.

plate. Figure 3a shows the values of c_3 deduced from equation (2) for several film thicknesses, as a function of temperature. These speeds are very low in comparison with ^4He , but they lie in the expected range for superfluid ^3He . At low temperatures, c_3 at first rises with δ ; but, when δ nears 170 nm, the speed begins to fall again. We believe that this behaviour arises from the competition between $\langle\rho_s\rangle$ rising with δ , and the van der Waals force F falling with δ .

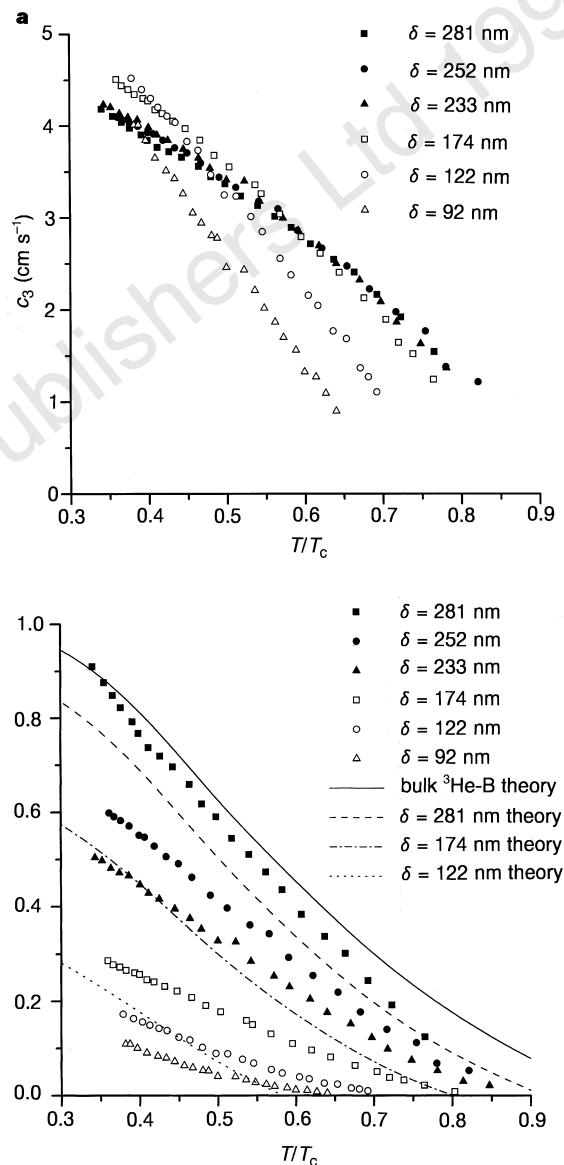


Figure 3 The speed of ^3He third sound and the associated superfluid density as a function of temperature. **a**, The speed of third-sound waves deduced from equation (2) for films with $\delta = 92, 122, 174, 233, 252$ and 281 nm, as a function of temperature. We show the value of c_3 as calculated from the frequency of the (0,3) mode using equation (2). The calculated values for c_3 using the frequencies of the other modes are in reasonable agreement with the data shown. The standard deviation for c_3 of any particular δ and T has a maximum value of 12%. For clarity, data is shown from only a subset of the measured films, but the speeds calculated for all the other films are also in good agreement with those shown. **b**, The average superfluid density of the film obtained from the speed of third sound and equation (1), as a function of temperature for films with $\delta = 92, 122, 174, 233, 252$ and 281 nm. The dashed lines represent the prediction of a model for a thin ^3He -B film with diffusive scattering of the Cooper pairs from the substrate and specular scattering from the surface of the film for three film thicknesses $\delta = 122, 174$ and 281 nm. The solid line is the bulk superfluid density of ^3He at zero pressure.

The superfluid fraction $\langle \rho_s \rangle / \rho$, contains fundamental information on the wave speed, the thermodynamics of the superfluid film, and the effects of disorder. To find $\langle \rho_s \rangle / \rho$ from equation (1), we need to know both δ and the van der Waals force F .

The value of F can be determined because the van der Waals potential $U(\delta)$ and the gravitational potential are equal at the surface of the film, that is

$$U(\delta) = gh(\delta) \quad (3)$$

where $h(\delta)$ is the height of the film above the bulk surface, and g is the acceleration due to gravity. This equilibrium occurs because the entire free surface of a superfluid is at a constant chemical potential²⁵. The film thickness is varied by introducing additional ^3He to reduce h . We find $U(\delta)$ by determining δ for various heights h , and fitting a smooth curve to our data for $h(\delta)$. Finally, we calculate the force from $F = -dU/d\delta = -gdh/d\delta$.

We determine δ from the capacitance between the substrate and the electrode above it. On the basis of comparison between $\delta(h)$ measurements in ^3He , and those in ^4He where both surfaces are covered, it seems that films of equal thickness reside on both electrodes. We assume that the electrodes are smooth and flat. Our measured $h(\delta)$ for ^3He can be expressed as $h = (\alpha/\delta)^n$ where $\alpha = (10.0 \pm 0.4) \times 10^{-9}$ (in units of metres^{4/3}) and $n = 3$. This is in reasonable agreement with previous measurements²⁰ which found $\alpha = (12.0 \pm 0.5) \times 10^{-9}$ metres^{4/3} and $n = 3$. The height h is measured using an annular capacitor which varies linearly with liquid height.

Substituting the measured values of δ , $F(\delta)$ and c_3 into equation (1), we find the values of $\langle \rho_s \rangle / \rho$ for the same films as in Fig. 3a. The results (Fig. 3b) indicate that as the films are made thicker, T_c^{film} rises (approaching the bulk value T_c). Also, for any given temperature, $\langle \rho_s \rangle / \rho$ rises with increasing film thickness. These observations are as expected when disorder in the substrate surface causes breaking of the ^3He Cooper pairs²². Suppression of superfluidity also occurs in the presence of strong disorder within the bulk liquid, for example, ^3He in aerogel²⁶. The dashed lines shown in Fig. 3b are theoretical predictions²⁷ for the $\langle \rho_s \rangle / \rho$ of thin superfluid ^3He -B films in the presence of disorder, and are in reasonable agreement with observations for our thickest films.

The quality of agreement between theory and observation for both the spectrum and $\langle \rho_s \rangle / \rho$, especially for the thickest films, is further evidence that the ^3He third sound is governed by equation (1). However, several unexpected observations have also been made with this probe. For films where $\delta < 3\xi$, $\langle \rho_s \rangle / \rho$ departs significantly from predictions and may indicate a change to an A-like phase, or the approach of true two-dimensionality in this superfluid²⁸. The complexity of the spectral features may also reflect unanticipated properties of this system. Although an isotropic superfluid film like ^4He should have no mixing of modes, ^3He is an anisotropic superfluid in which several sources of coupling might be possible, including textural effects, and a Hall effect phenomenon²⁹ due to the ordering of the angular momentum of the Cooper pairs in the A-phase. These issues will be addressed in future work.

The observation of third sound in superfluid ^3He opens the door to the study of a number of important phenomena for which, until now, no experimental technique was available. These include the limit on how thin a ^3He film can be made and still remain superfluid, the transition to two-dimensional superfluidity possibly via a Kosterlitz–Thouless vortex unbinding phase transition²⁸, the superfluid analogue of the quantum Hall effect²⁹, and the potential for two-dimensional ^3He superfluidity in submonolayer coverages via a new mechanism³⁰. □

Received 1 July; accepted 29 September 1998.

- Landau, L. D. & Lifshitz, E. M. *Fluid Mechanics* 36–44 (Pergamon, Oxford, 1959).
- Bildsten, L. & Cutler, C. Nonradial oscillations in neutron star oceans: a source of quasi-periodic x-ray oscillations? *Astrophys. J.* **449**, 800–812 (1995).

- Whitaker, D. L., Weilart, M. A., Vincente, C. L., Maris, H. J. & Seidel, G. M. Oscillations of charged helium II drops. *J. Low Temp. Phys.* **110**, 173–178 (1998).
- Tilley, D. R. & Tilley, J. *Superfluidity and Superconductivity* 77–83 (Hilger, Bristol, 1986).
- Atkins, K. R. Third and fourth sound in liquid helium II. *Phys. Rev.* **113**, 962–965 (1959).
- Everitt, C. W. F., Atkins, K. R. & Denenstein, A. Third sound in liquid helium films. *Phys. Rev.* **136**, A1494–A1499 (1964).
- Bergman, D. Hydrodynamics and third sound in thin HeII films. *Phys. Rev.* **188**, 370–384 (1969).
- Rudnick, I. Critical surface density of the superfluid component in ^4He films. *Phys. Rev. Lett.* **40**, 1454–1455 (1978).
- Ellis, F. M., Hallock, R. B., Miller, M. D. & Guyer, R. A. Phase separation in films of ^3He - ^4He mixtures. *Phys. Rev. Lett.* **46**, 1461–1464 (1981).
- Ketola, K. S., Wang, S. & Hallock, R. B. Anomalous wetting of helium on cesium. *Phys. Rev. Lett.* **68**, 201–204 (1992).
- Ellis, F. M. & Li, L. Quantum swirling of superfluid helium films. *Phys. Rev. Lett.* **71**, 1577–1580 (1993).
- Sheldon, P. A. & Hallock, R. B. Third sound and energetics in ^3He - ^4He mixture films. *Phys. Rev. B* **50**, 16082–16085 (1994).
- Brisson, J.-G., Mester, J. C. & Silveira, I. F. Third sound of helium on a hydrogen substrate. *Phys. Rev. B* **44**, 12453–12462 (1991).
- Ekholm, D. T. & Hallock, R. B. Studies of the decay of persistent currents in unsaturated films of superfluid ^4He . *Phys. Rev. B* **21**, 3902–3912 (1980).
- Pickar, K. A. & Atkins, K. R. Critical velocity of a superflowing liquid-helium film using third sound. *Phys. Rev.* **178**, 389–399 (1969).
- Cho, H. & Williams, G. A. Vortex core size in submonolayer superfluid ^4He films. *Phys. Rev. Lett.* **75**, 1562–1565 (1995).
- Eggenkamp, M. E. W. *et al.* On the possibility of observing third sound on ^3He . *J. Low Temp. Phys.* **110**, 299–304 (1998).
- Schechter, A. M. R., Simmonds, R. W. & Davis, J. C. Capacitive generation and detection of third sound resonances in saturated superfluid ^4He films. *J. Low Temp. Phys.* **110**, 603–608 (1998).
- Sachrajda, A., Harris-Lowe, R. F., Harrison, J. P., Turkington, R. R. & Daunt, J. G. ^3He film flow: two-dimensional superfluidity. *Phys. Rev. Lett.* **55**, 1602–1605 (1985).
- Daunt, J. G. *et al.* Critical temperature and critical current of thin-film superfluid ^3He . *J. Low Temp. Phys.* **70**, 547–568 (1988).
- Davis, J. C., Amar, A., Pekola, J. P. & Packard, R. E. Superfluidity of ^3He films. *Phys. Rev. Lett.* **60**, 302–304 (1988).
- Xu, J. & Crooker, B. C. Very thin films of ^3He : a new phase? *Phys. Rev. Lett.* **65**, 3005–3008 (1990).
- Freeman, M. R., Germain, R. S., Thuneberg, E. V. & Richardson, R. C. Size effects in thin films of superfluid ^3He . *Phys. Rev. Lett.* **60**, 596–599 (1988).
- Jacobsen, K. W. & Smith, H. Critical current for the flow of superfluid ^3He in a confined geometry. *J. Low Temp. Phys.* **67**, 83–89 (1987).
- Tilley, D. R. & Tilley, J. *Superfluidity and Superconductivity* 92–94 (Hilger, Bristol, 1986).
- Porto, J. V. & Parpia, J. M. Superfluid ^3He in aerogel. *Phys. Rev. Lett.* **74**, 4667–4670 (1995).
- Freeman, M. R. *Size Effects in ^3He Films*, Thesis, Cornell Univ. (1988).
- Stein, D. L. & Cross, M. C. Phase transitions in two dimensional superfluid ^3He . *Phys. Rev. Lett.* **42**, 504–507 (1979).
- Volovik, G. E. An analog of the quantum Hall Effect in a superfluid ^3He film. *Sov. Phys. JETP* **67**, 1804–1811 (1988).
- Kurihara, S. Singlet superfluidity in ^3He film on ^4He film. *J. Phys. Soc. Jpn* **52**, 1311–1316 (1983).

Acknowledgements. We thank L. Bildsten, R. B. Hallock, D.-H. Lee, J. M. Goodkind, O. Ishikawa, J. Harrison and G. Williams for discussions. This work was supported by the NSF, the ONR and by the Packard Foundation.

Correspondence and requests for information should be addressed to J.C.D. (e-mail: jcdavis@socrates.berkeley.edu); see also <http://socrates.berkeley.edu/~davisgrp>.

Monochromatic electron emission from the macroscopic quantum state of a superconductor

K. Nagaoka*, T. Yamashita*, S. Uchiyama*, M. Yamada*, H. Fujii* & C. Oshima*†

* Department of Applied Physics, Waseda University, 3-4-1, Okubo, Shinjyuku-ku, Tokyo 169-8555, Japan

† Kagami Memorial Laboratory for Material Science and Technologies, Waseda University, 2-8-26, Nishiwaseda, Shinjyuku, Tokyo 169-0051, Japan

The ground state of a superconductor is a macroscopic quantum state that can extend coherently over substantial distances¹. As a result, electrons tunnelling from two different points (separated by macroscopic length) on the surface of a superconductor remain coherent in phase and so are able to interfere: this property forms the basis of superconducting quantum interference devices (SQUIDS). Another characteristic of electrons tunnelling from a superconductor is that they are monochromatic, their energy being determined by the ground-state energy of the superconducting state. Monochromatic electrons have been observed

tunnelling from a superconductor to a normal metal², and the resulting currents have been used to probe the dynamics of atoms and molecules at interfaces³. Here we report the results of field-emission experiments that confirm the prediction⁴ that monochromatic electrons can similarly be emitted from a superconductor into vacuum. Monochromatic emissions of this type might find application as the sources in a range of electron-based spectroscopies.

Monochromatic electron beams emitted into vacuum from the surface of a superconductor have been predicted theoretically⁴, but have hitherto eluded experimental confirmation in spite of considerable efforts^{5–7}. Over the past several years, we have improved the performance of high-resolution electron spectroscopy for field emission (FE), and the experimental technique of extremely high vacuum^{8–11}. By combining these techniques with a cryo-FE gun in this work, we have detected (to our knowledge, for the first time) the emission of a monochromatic beam into vacuum from the Bose–Einstein condensed state of Cooper pairs in a niobium superconductor.

Figure 1 shows a schematic diagram of the apparatus used in this experiment. The cryo-FE gun is isolated thermally from room-temperature regions; to reduce conducted heat flow, all the electric leads connected to the FE gun pass through the liquid-helium tank, and thermal radiation was reduced by cooling to ~ 4.2 K the cryo-walls and electrodes. The Nb tip was mounted on a tip holder made of annealed pure Cu (99.9%) with high thermal conductivity, of which one edge directly touches the liquid helium. There are two small holes in the cryo-walls. One is the entrance of the electron analyser located just below the tip. Through another hole, the FE patterns are observed by using an optical fibre, which was removed from the cryo-gun during low-temperature experiments. The temperature at the position indicated by the letter T in Fig. 1 was measured with a chromel–Au/Fe thermocouple, and it was calibrated with the superconducting transition temperature of a Nb

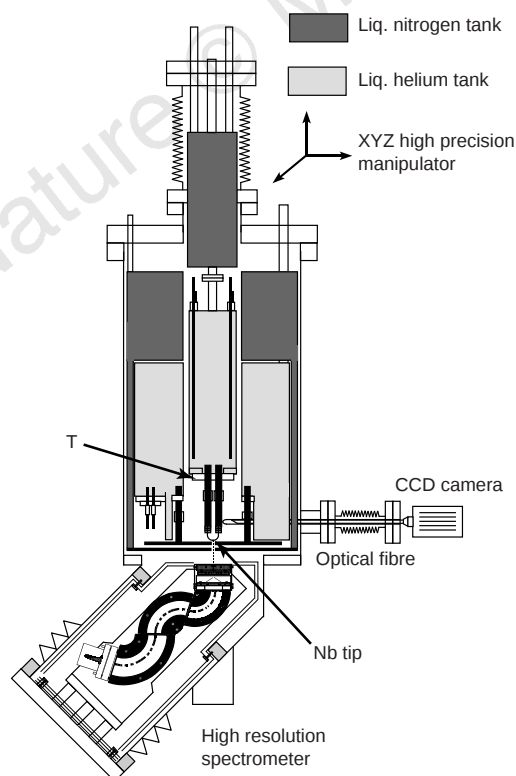


Figure 1 A schematic diagram of the apparatus used in this experiment. It consists of a cryo-FE electron gun, a system to observe the FE pattern, and a high-resolution electron energy analyser. See text for details of position marked T.

wire at the emitter position. The temperature difference between the measuring point and the emitter position was less than ± 1 K. The specimens are polycrystalline Nb tips with a radius of ~ 100 nm fabricated by electrochemical polishing. The tip shape was evaluated from the FE pattern. The tip surface was finally cleaned by repeated field evaporation in extremely high vacuum. Typical emission currents measured in this experiment were less than $\sim 10^{-7}$ A under extracting voltages of ~ 1.2 kV, and the base pressure was less than 1×10^{-9} Pa. The energy resolution of the whole spectrometer was estimated to be less than 1 meV.

Figure 2 shows typical spectra of the FE beam from a Nb tip at 25 K and 4.2 K, which are respectively above and below the superconducting transition temperature of $T_c = 9.2$ K. The electron intensity was plotted on a logarithmic scale against the energy referred to the Fermi level (E_F). The filled circles represent experimental points, and the dashed curves are calculated from Fowler–Nordheim (FN) theory¹². The curve fitting was done by adjusting three parameters; work function (Φ), electric field (F) and tip temperature (T). For fitting the gradient of the low-energy slope, at first we adjusted Φ and F (ensuring that these parameters were consistent with other data such as FN plots and tip radius), and then calculated the high-energy slope with the observed temperature T . In this procedure, we have noticed the necessity of an additional broadening of ~ 25 meV when using reasonable parameters for all observed spectra; that is, we could not reproduce the high-energy slope when using only the observed temperature. Three parameters used in Fig. 2 are 4.36 eV for Φ , 4.2 K (25 K) for T , and 3.4 (3.9) V nm^{-1} for F . Except for the tiny difference originating from surface band structures⁹, and the hot-hole injection generated by FE¹⁰, the curve at 25 K was successfully reproduced from the FN distribution with an additional broadening of 25 meV. In contrast, it should be noticed that a new sharp peak appears just at E_F in the spectrum at 4.2 K.

Figure 3 shows the change in the spectra with temperature around T_c . Only the narrow-energy region around E_F is shown in the spectra. An extra sharp peak appears around E_F below T_c , and the intensity increases with decreasing temperature. The new peak intensity is plotted against the temperature normalized by T_c in Fig. 3 inset. The intensity change agrees with the solid curve, which represents the temperature dependence of the order parameter calculated on the basis of BCS theory, namely the gap parameter, which is proportional to the square root of Cooper-pair density. This is experimental evidence that the electrons in the extra

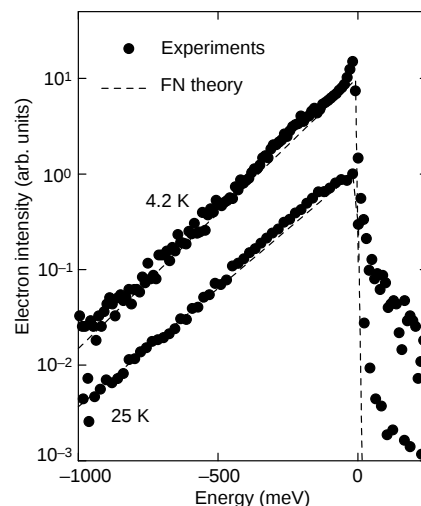


Figure 2 Typical FE spectra from the Nb tip at 4.2 K and 25 K, which are respectively below and above T_c . The dashed curves are calculated on the basis of the FN theory. A sharp peak appears at E_F in the spectrum at 4.2 K.

peak are correlated with the superconducting states, namely the Bose–Einstein condensed states of the Cooper pairs.

In this experiment, furthermore, we have found strong correlation between the sharp peak and surface contamination. In general, all molecules except He and H₂ in the residual gases condense onto the tip surface at 4.2 K, resulting in the formation of a thick physisorbed layer. When the surface was covered with the thick films of adsorbates, the extra peak did not appear in the spectra. Only repeated surface cleaning by field evaporation made it possible to reproduce the extra peak. Here, we applied the high voltage ~6 kV, which is one order of magnitude larger than those of normal FE. Even at a vacuum pressure of 10⁻¹⁰ Pa, the peak intensity tends to reduce with time, and finally disappears after a long time. As the coherence length of ~38 nm in the Nb superconductor¹ is shorter than the tip radius of ~100 nm, the superconducting phase may occupy the volume inside the tip apex, but not exist just at the surface; hence, the tail of the wavefunction of the superconducting phase can penetrate into vacuum through the clean tip apex without contaminated layers, but not through the contaminated layers. This is presumably the reason why the extra peak was observed only for clean tips.

In the pioneering work by Gomer and Hulm for a Ta tip⁵ and Klein and Leder for a Nb tip⁶, no significant change in emission current related to the superconducting transition was detected around T_c . The difference from the present work was the absence of surface cleaning; the thick physisorbed layers prevented electron emission from the superconducting region.

The observed energy spread of the extra peak (full-width at h. l. maximum) is ~20 meV, which is much broader than that predicted by theory (<0.1 meV, on the basis of the BCS theory for highly pure Nb metal at 0 K; ref. 4). As we mentioned in discussing Fig. 2, the same broadening appeared in the spectra above T_c . The curve fitting required the other broadening mechanism of ~25 meV in addition to the temperature broadening. Why does it broaden? The following are possible causes:

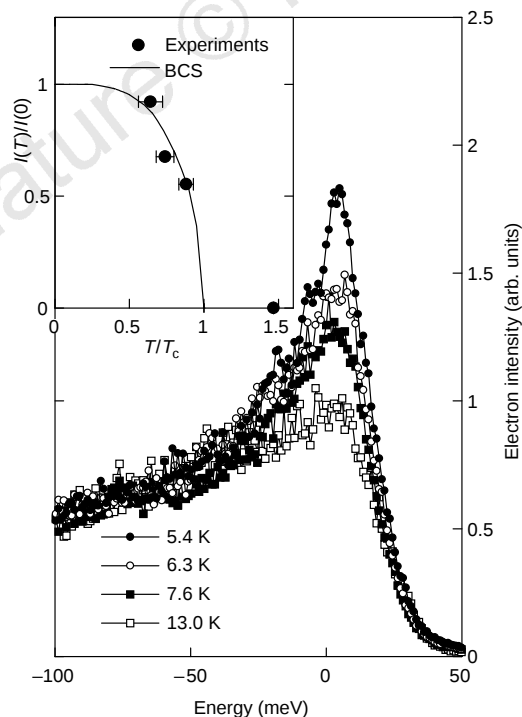


Figure 3 The observed FE spectra from the Nb tip at various temperatures around T_c (=9.2 K). Inset, the intensity of the sharp peak plotted against the temperature normalized by T_c . The solid curve represents the temperature dependence of the order parameter in BCS theory.

(1) Coulomb interaction between the emitted electrons in vacuum, so-called Boers effects, in which high-density electrons interact with each other, resulting in the energy spread. In this experiment, however, no change in the energy spread was detected in spectra measured at different currents from 10⁻⁷ to 10⁻⁸ A, corresponding to current densities of 20 to 2 A cm⁻², which is consistent with Bell and Swanson's results¹³; the total current of 10⁻⁷ A used in this experiment is three orders of magnitude lower than the practical current density at which the Boers effect appears. In this case, therefore, the Coulomb interaction effect seems to be negligibly small.

(2) An inelastic tunnelling effect in which the excitations of surface atomic vibration and surface phonons produce an energy spread. However, previous work concerning inelastic tunnelling indicated an extremely small cross-section of the excitations as compared with the observed extra peak^{9,14}. It is also consistent with many data from inelastic tunnelling spectroscopy³.

(3) Lifetime broadening, discussed by Gadzuk⁴. The electrons occupying certain states in metals possess their own lifetime determined mainly by imperfections of the lattice such as impurities and defects. The electric residual resistivity of the Nb wire without special purification is ~1.3 × 10⁻⁷ Ω cm (ref. 15), which corresponds to a relaxation time of 1.0 × 10⁻¹³ s, and also to lifetime broadening of 6.3 meV from the uncertainty principle. The expected value is of the same order as the observed broadening.

We would like to point out the uncertainty of the energy resolution of the spectrometer. In our previous work⁸, we estimated this quantity to be <1 meV from the beam spread inside the third analyser. In this experiment, however, the spectra were obtained by retarding all the voltages supplying the spectrometer, which differs from the procedure followed for estimation of the energy resolution. Further work seems to be required to clarify the origin of this broadening.

We have detected a characteristic feature in the FE spectra from a superconducting region of a niobium tip. The intensity of the extra peak increases with decreasing temperature below T_c , with a similar temperature dependence to that of the order parameter. The observed energy spread of ~20 meV should be able to be reduced drastically either by improving the energy resolution of the electron energy analyser or by using a niobium specimen without imperfections. □

Received 23 March; accepted 1 September 1998.

- Kittel, C. *Introduction to Solid State Physics* 7th edn (Wiley, New York, 1996).
- Glauber, I. Electron tunneling between two superconductors. *Phys. Rev. Lett.* **5**, 464–466 (1960).
- Poppe, U. & Wuhl, H. Tunneling spectroscopy on the superconducting chevre-phase compounds Cu_{1.8}Mo_{0.8}S₈ and PbMo_{0.8}S₈. *J. Low Temp. Phys.* **43**, 371–382 (1981).
- Gadzuk, J. W. Many-body tunneling-theory approach to field emission of electrons from solids. *Surf. Sci.* **15**, 466–482 (1969).
- Gomer, R. & Hulm, J. K. Field emission from tantalum in the normal and superconducting state. *J. Chem. Phys.* **20**, 1500–1502 (1953).
- Klein, R. & Leder, L. B. Field emission from niobium in the normal and superconducting states. *Phys. Rev.* **124**, 1050–1052 (1961).
- Bergret, H., Septier, A. & Drechsler, M. Nottingham effect of a superconducting metal. *Phys. Rev. B* **31**, 149–153 (1985).
- Oshima, C. An apparatus for high resolution field emission spectroscopy. *Adv. Colloid Interface Sci.* **71/72**, 353–369 (1997).
- Nagaoka, K. et al. High resolution field emission spectra from niobium (100) tip surfaces. *Sci. Rep. RITU A* **44**, 257–260 (1997).
- Ogawa, H. et al. Energy spectra of field emission electrons from a W (310) tip. *Surf. Sci.* **357–358**, 371–375 (1996).
- Nagaoka, K. et al. Angular resolved field emission spectra from Nb tips. *Appl. Surf. Sci.* **130–132**, 512–517 (1998).
- Young, R. D. Theoretical total-energy distribution of field-emitted electrons. *Phys. Rev.* **113**, 110–114 (1959).
- Bell, A. E. & Swanson, L. W. Total energy distributions of field-emitted electrons at high current density. *Phys. Rev. B* **19**, 3353–3364 (1979).
- Plummer, E. W. & Bell, A. E. Field emission energy distributions of hydrogen and deuterium on the (100) and (110) planes of tungsten. *J. Vac. Sci. Technol.* **9**, 583–590 (1971).
- Takeuchi, N. & Inaba, S. Refinement of Ta and Nb under UHV conditions. *Solid Stat. Phys.* **15**, 113–115 (1980). (In Japanese.)

Acknowledgements. We thank N. Osakabe, H. Bergret and T. Sakurai for providing the experimental techniques for the cryo-electron gun, and S. Kurihara for comments concerning the interpretation. This work was supported partly by Special Coordination Funds for Promoting Science and Technology and partly by Research for the Future in JSPS.

Correspondence and requests for materials should be addressed to C.O. (e-mail: coshima@mn.waseda.ac.jp).

Quantification of decadal anthropogenic CO₂ uptake in the ocean based on dissolved inorganic carbon measurements

Tsung-Hung Peng*, Rik Wanninkhof*, John L. Bullister†, Richard A. Feely† & Taro Takahashi‡

* Atlantic Oceanographic and Meteorological Laboratory, NOAA, Miami, Florida 33149, USA

† Pacific Marine Environmental Laboratory, NOAA, Seattle, Washington 98115, USA

‡ Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York 10964, USA

About half of the 'anthropogenic' CO₂ emitted to the atmosphere is taken up by the oceans and terrestrial biosphere¹, and the amount sequestered by the ocean is generally estimated using numerical ocean carbon-cycle models². But these models often differ markedly³, resulting in different estimated spatial and temporal patterns and magnitudes of uptake. Because of its importance climatically, the CO₂ flux needs to be verified using field measurements. Accurate estimates of CO₂ uptake have been difficult to obtain, however, as the annual increase of dissolved inorganic carbon (DIC) concentration in surface water due to anthropogenic input is ~0.05% of the total DIC, an order of magnitude lower than past measurement precision. Early measurement-based estimates^{4,5} of total anthropogenic CO₂ inventory in the ocean have recently been improved on^{6,7}, and new approaches have been proposed for determining changes in ocean DIC concentration over one to two decades^{8,9}. Here we use recent improvements in DIC measurement techniques to determine changes in DIC concentrations between 1978 and 1995 in the Indian Ocean. Our method subtracts decadal-scale natural variability, enabling the ocean anthropogenic CO₂ increase in this region over the 17-year period to be determined. The calculated uncertainties and known measurement capabilities allow us to define the minimum sampling strategies that will be required to quantify the regional and global anthropogenic CO₂ oceanic uptake over future decades.

To derive temporal changes in oceanic CO₂ uptake, the most direct method is a straight comparison of measurements of DIC at two different times. Changes in DIC derived from such a comparison reflect the anthropogenic input over this time period, assuming an insignificant change in water masses during the measurement period. Small changes due to natural variability can be accounted for by covariance of carbon with apparent oxygen utilization (AOU). To enable such a comparison to be made, we need to have a data set of reasonable quality taken in the past to serve as a

reference. A global survey of geochemical properties was conducted between 1973 and 1978 during the Geochemical Ocean Sections Study (GEOSECS) program. New measurements have been made more recently during global CO₂ survey cruises. This data set provides an opportunity to examine the DIC increase over almost a two-decade period.

The comparison in the Indian Ocean is used to illustrate the technique and uncertainties associated with removing the decadal natural variations. Six stations were occupied along 80° E from 5° N to 20° S during the GEOSECS expedition in April 1978. This cruise track was re-occupied during a CO₂ survey cruise (I8NR) in 1995. (See Methods for a description of measurements, methodology and precision.)

Because no common references among these sets of measurements are available, data from the two observational periods have to be converted to a common scale. Most of the water column below 2,000 m depth in this region of the Indian Ocean has had negligible contact with the atmosphere during the past 40 years¹⁰, as suggested by the near absence of chlorofluorocarbons. These waters respond very slowly on decadal timescales to changes in surface conditions. Thus, the chemical properties, such as DIC, alkalinity (Alk) and dissolved O₂, should remain the same in deep water if there are no systematic errors between the cruises and if water masses remain the same.

We use the assumed invariance of the properties of deep water to correct the GEOSECS data. Deep-water concentrations of Alk and O₂ are the same for both sets of measurements (Table 1). However, a systematic difference between the GEOSECS and I8NR DIC data is seen. All GEOSECS values are $21.3 \pm 3.4 \mu\text{mol kg}^{-1}$ higher than those of the more recent I8NR cruise (Table 1). Part of the error is caused by the difference in analysis techniques for DIC measurement. The more recent coulometric method determines the amount of CO₂ gas liberated from acidified seawater samples. The GEOSECS data were obtained using potentiometric titration of carbonic acid in sea water, and show a positive bias due to combinations of various shortcomings, including non-ideal behaviour of pH electrodes, imperfection of the ionic model used to interpret the titration data, and the possible presence of various organic acids.

To minimize the influence of variations in thermocline depth over time, the DIC comparison is made along the same isopycnal horizon. We interpolate the observed properties onto a few designated isopycnal surfaces. To avoid changes in the surface ocean caused mainly by seasonal variations, we chose the isopycnal horizons in the upper thermocline. In the Indian Ocean, these waters are represented by potential density σ_θ values ranging from 26.6 to 27.2, which have a depth range of 300–1,000 m. At these depths, the variations caused by the seasonal effect are expected to be minimal. The outcrop area for the isopycnals in question is south of 35° S.

To reduce artefacts due to small natural variations in biological respiration between the two different periods, we need to correct for the amount of DIC derived from respiration products. This is done by using the AOU, which is defined as the difference between the observed dissolved oxygen concentration and the saturated oxygen

Table 1 Changes of chemical properties on isopycnals between two surveys

σ_θ	ΔDIC	s.d.*	ΔO_2	s.d.	ΔAOU	s.d.	ΔAlk	s.d.	ΔSal	s.d.	N
26.6	8.7	12.5	2.9	12.1	-2.1	12.7	2.2	2.4	-0.019	0.053	5
26.7	17.9	8.3	-4.4	6.4	6.6	5.5	-0.8	2.6	-0.067	0.053	4
26.8	6.3	10.5	3.2	17.2	-0.4	17.6	-0.6	3.2	-0.016	0.045	6
26.9	15.2	12.3	-5.5	19.8	6.7	19.3	3.3	5.3	-0.012	0.026	5
27.0	15.3	6.8	-11.7	7.8	11.7	8.2	5.4	3.4	0.005	0.022	6
27.2	5.1	6.8	-4.9	7.3	5.2	6.8	6.2	3.3	0.014	0.025	6
Deep	-21.3	3.4	-1.0	2.4	-	-	-1.1	3.2	-	-	6

These are mean differences between I8NR and GEOSECS for DIC, O₂, AOU, Alk, and Sal on isopycnals in the latitude zone between 5° N and 20° S. All DIC, O₂ and Alk values are corrected for deepwater offset given in the last line. The difference Δ equals I8NR minus GEOSECS.

* s.d., Standard deviation.

concentration at the potential temperature and salinity of the sample. The effect of changes in dissolution of calcium carbonate tests at the two different times is corrected by using Alk data. To eliminate changes associated with variations in salinity, the corrected DIC is normalized to a salinity of 35.0 (given as DIC_n ; see Methods).

The I8NR survey shows a significant increase in DIC_n compared with the GEOSECS survey (Fig. 1). Furthermore, the increases along the lighter isopycnals are greater than the denser isopycnals. The two curves are essentially indistinguishable from each other when the density reaches 27.2. This trend in difference between two curves is consistent with the prediction that the shallower (or lighter) waters should contain more anthropogenic CO_2 , as these waters should have interacted more recently with the atmosphere containing greater concentrations of CO_2 . In addition, the magnitude of changes is significantly larger in the latitude zone between 10° S and 20° S for all density horizons analysed here (Table 2), showing that more anthropogenic carbon is taken up in the temperate zone than in the equatorial zone during the period between the two surveys, which is consistent with bomb ^{14}C data¹¹. Results of chlorofluorocarbon measurements during the I8NR cruise also show the highest concentrations at the shallower isopycnals, and a pronounced north–south gradient, with higher concentrations (by a factor of ~5) found to the south, which is also in accordance with the pattern of DIC_n increase over time.

For the difference in DIC_n observed between two time periods, it is reasonable to assume that the mixing processes remain unchanged and the difference is mainly caused by the invasion of anthropogenic CO_2 . T – S profiles do not indicate a significant change in water mass between the GEOSECS and I8NR surveys. Thus, the difference in DIC_n is assumed to represent the anthropogenic CO_2 signal. The mean difference of DIC between the two cruises on the same isopycnal surface is obtained by integrating the area between the two curves shown in Fig. 1. The DIC increases range from 6 to 25 $\mu\text{mol kg}^{-1}$ for the region between 20° S and 10° S, and 1 to 5 $\mu\text{mol kg}^{-1}$ for the region between 10° S and 5° N (Table 2). For the region as a whole, the 26.6 isopycnal has a DIC increase of the order of 12 $\mu\text{mol kg}^{-1}$, and this excess DIC decreases with the increasing density (Fig. 2). There is no detectable excess DIC at an isopycnal of 27.2. Based on the precision of DIC measurements from both cruises, and six regions for comparison, we estimate the random error for the mean difference to be about $\pm 4.5 \mu\text{mol kg}^{-1}$.

To verify that the changes can be predominantly attributed to the anthropogenic signal, the differences in O_2 , AOU, Alk, salinity and DIC between GEOSECS and I8NR for the selected isopycnals are shown in Table 1. The interpolated values on these isopycnals in 1978 are compared with those from the stations with the same geographical location re-occupied in 1995. The differences for the isopycnal surfaces are negligible for salinity but can be significant for AOU and Alk. As the AOU and Alk are used to distinguish

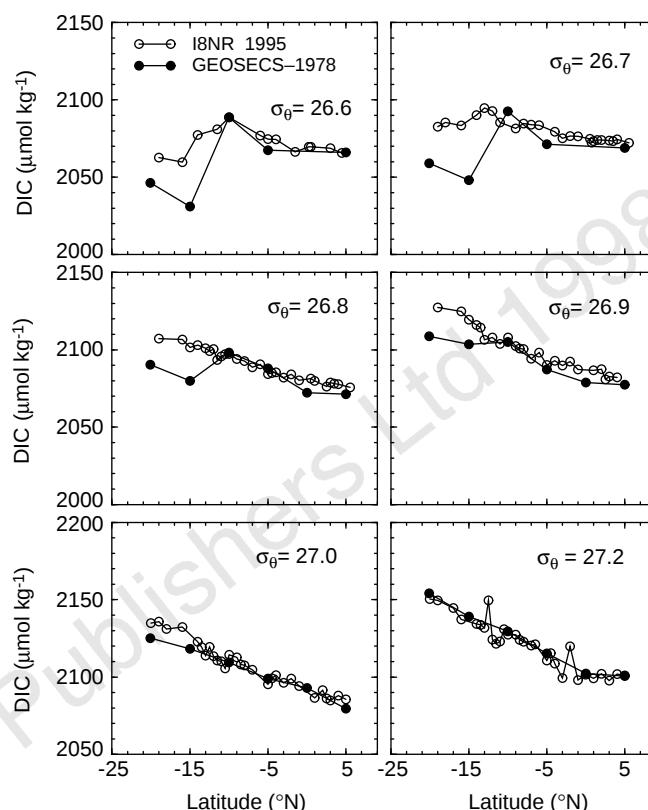


Figure 1 DIC increase in the upper thermocline. A comparison is shown of salinity-normalized DIC, corrected for AOU and Alk changes over time, along an isopycnal surface in the upper thermocline of the Indian Ocean along 80° E longitude for a latitude zone ranging from 20° S to 5° N at two different times: GEOSECS in 1978 (filled circles) and I8NR in 1995 (empty circles). The different panels represent six isopycnal surfaces at σ_θ values of 26.6, 26.7, 26.8, 26.9, 27.0 and 27.2. The I8NR is a CO_2 survey cruise operated by NOAA Ocean–Atmosphere Carbon Exchange Study program together with the Joint Global Ocean Flux Study (JGOFS) and World Ocean Circulation Experiment (WOCE) repeat hydrographic survey.

between natural variations in DIC and anthropogenic loading (see Methods), differences over time will contribute to the uncertainty in DIC changes. The published range of Redfield ratios¹² between DIC and O_2 is 0.69 to 0.81, which translates into an uncertainty of 0.1 to 1.2 $\mu\text{mol kg}^{-1}$ in the natural variations of DIC on the isopycnal surfaces investigated. This is a small effect as this method is based on differences of two data sets along isopycnals where anthropogenic perturbation is far greater than natural variability.

Table 2 Anthropogenic CO_2 uptake in the ocean

σ_θ	DIC_0	DIC_1	DIC_2	H_1	H_2	SI_1	SI_2	Inv_1	Inv_2
		$(\mu\text{mol kg}^{-1})$		(m)		(mol m^{-2})		(10^{13} mol)	
26.6	10.5	21.4	3.2	337	255	7.40	0.84	7.55	1.10
26.7	12.3	25.1	3.8	52	45	1.34	0.18	1.37	0.24
26.8	8.1	15.9	2.9	58	67	0.95	0.20	0.97	0.26
26.9	8.3	12.5	5.5	70	84	0.90	0.47	0.92	0.62
27.0	3.3	6.2	1.4	72	86	0.46	0.12	0.47	0.16
Total				589	537	11.05	1.81	11.28	2.38

Mean DIC inventory change:

20° S to 10° S = 1.1 $\mu\text{mol kg}^{-1} \text{ yr}^{-1}$

10° S to 5° N = 0.2 $\mu\text{mol kg}^{-1} \text{ yr}^{-1}$

Excess CO_2 inventory: $1.64 \times 10^{15} \text{ g C}$

These are specific and total anthropogenic CO_2 inventory in the Indian Ocean along 80° E from 20° S to 5° N during a period from 1978 to 1995. DIC_0 : For the whole region 20° S to 5° N (area, $2.33 \times 10^{13} \text{ m}^2$); DIC_1 : 20° S to 10° S region (area, $1.02 \times 10^{13} \text{ m}^2$); DIC_2 : 10° S to 5° N region (area, $1.31 \times 10^{13} \text{ m}^2$); H_1 and H_2 : mean thickness of isopycnals (m); SI_1 and SI_2 : specific inventory (mol m^{-2}); Inv_1 and Inv_2 : total inventory ($\times 10^{13} \text{ mol}$).

Part of the reason for the variations of estimated DIC increases along isopycnals (Fig. 2a) is probably that the GEOSECS data are less precise. In future, better data can be compared with the high-quality I8NR data as a baseline for detecting the increase in the CO₂ inventory over the next decade. Current estimates of accuracy of DIC are $\pm 1.5 \mu\text{mol kg}^{-1}$, whereas the precision for Alk and O₂ is 2 and $0.2 \mu\text{mol kg}^{-1}$, respectively, making detection of the anthropogenic signal in the ocean over time more robust. Moreover, anthropogenic loading is accelerating¹, so the signal in the ocean is increasing at a greater rate as well.

The magnitude of DIC increases can be compared with atmospheric CO₂ increases. Although the ocean water of outcrop region for these isopycnals may not be at equilibrium with atmospheric CO₂ content, in order to estimate the maximum possible CO₂ signal we assume that the ocean uptake keeps up with atmospheric increases. As shown in Fig. 2b, for a time period of 17 years between GEOSECS and I8NR, the DIC increase could be in the range of 5 to $17 \mu\text{mol kg}^{-1}$, depending on the rate of atmospheric CO₂ increase and on when the water was ventilated. Also shown is the predicted increase if the ocean should follow the observed atmospheric increase¹. The magnitude of the DIC increase estimated here is consistent with these predictions.

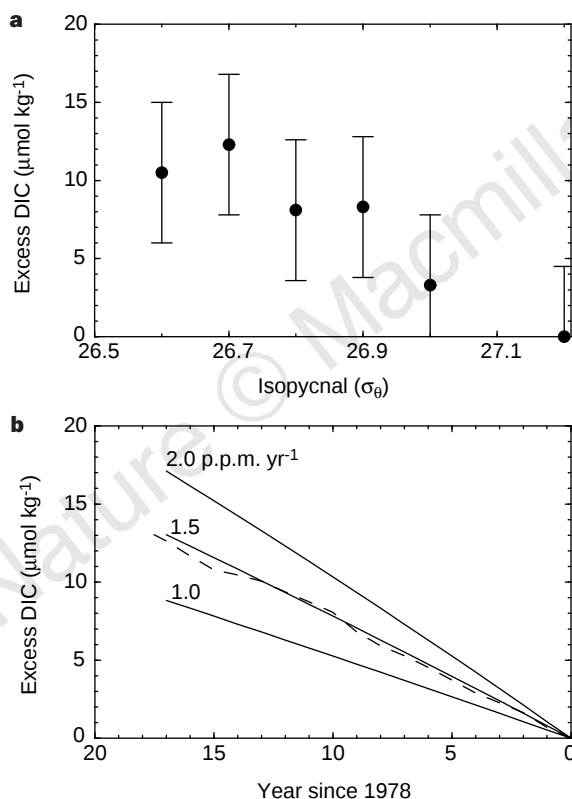


Figure 2 Anthropogenic CO₂ signal in the ocean. **a**, The mean difference of DIC concentrations between GEOSECS and I8NR is plotted against density surfaces. This is the anthropogenic CO₂ signal in the CO₂ survey area of the Indian Ocean along 80° E. The error of $4.5 \mu\text{mol kg}^{-1}$ is estimated based on DIC random error of $11 \mu\text{mol kg}^{-1}$ during GEOSECS times, and of $1.4 \mu\text{mol kg}^{-1}$ during the I8NR, using 6 regions for making the estimate of the mean signal strength for the particular isopycnal surface. **b**, Excess DIC since 1978 in equilibrium with the stated atmospheric CO₂ growth rate is plotted against the elapsed time. For comparison with the observed oceanic anthropogenic CO₂ signal in **a**, the time axis is reversed so that the water equilibrated further back in time is equivalent to water with higher density. The excess DIC is the difference in DIC at equilibrium with atmospheric CO₂ content between the year 1978 and subsequent years. The dotted line is based on the observed record of atmospheric CO₂ content at the Mauna Loa Observatory¹.

The oceanic invasion of anthropogenic CO₂ in the study region can be estimated by integrating the DIC changes between these two cruises over the water column (Table 2). The mean depth for the density surface of 26.6 is obtained by assuming that the excess DIC on this isopycnal extended all the way to the surface mixed layer. This assumption leads to a lower limit of the inventory because the excess CO₂ above 26.6 is higher than that in this isopycnal. The results of an annual mean, three-dimensional ocean model^{13,14} of the DIC increase between 1978 and 1995 in our study area agree well with our observations for the entire transect, with specific excess CO₂ inventories of 5.5 ± 2.2 and 5.8 mol m^{-2} , respectively, but differ in location of storage. The mean water column excess DIC for the region between 22.2°S and 8.9°S in the model is $7.1 \pm 1.6 \text{ mol m}^{-2}$, which is $\sim 55\%$ less than our estimate of 11.1 mol m^{-2} for the region between 20°S and 10°S (Table 2). For the region between 8.9°S and 4.5°N, the model gives $3.9 \pm 0.2 \text{ mol m}^{-2}$; our estimate is 1.8 mol m^{-2} for the region between 10°S and 5°N. This comparison indicates that the model does not store enough CO₂ in the temperate region, and may over predict storage in the equatorial region. With an improved direct estimate, such spatial variations in temporal oceanic CO₂ uptake in the future can be used for evaluating the model performance and so for improving model construction. The rate of DIC increase during this time period is estimated to be $0.2 \mu\text{mol kg}^{-1} \text{ yr}^{-1}$ for the equatorial region, and $1.1 \mu\text{mol kg}^{-1} \text{ yr}^{-1}$ for the temperate region. The mean increase for the whole region is estimated to be $0.5 \mu\text{mol kg}^{-1} \text{ yr}^{-1}$. The total accumulated anthropogenic CO₂ for the whole width of the Indian Ocean between 20°S and 5°N is estimated to be $1.64 \times 10^{15} \text{ g C}$ during this period. This is compatible with the Princeton model estimate of $1.47 \times 10^{15} \text{ g C}$ for the same area in this period¹⁴.

Our study demonstrates that the anthropogenic CO₂ signal can be detected on a decadal timescale by high-precision DIC measurement. By using current high-quality data as a baseline, surveys over the next decade should provide a clearer picture of the anthropogenic CO₂ loading in the ocean, not least because of the continued exponential rise of atmospheric CO₂ levels. Sampling density in a future repeat survey can be evaluated. By reducing I8NR sampling stations to only six, corresponding to a station spacing of about 5 degrees, the estimated CO₂ uptake is similar to within 5% to the results presented here. The detection of the anticipated increase in DIC of over $10 \mu\text{mol kg}^{-1}$ in the next ten years in the upper thermocline should be straightforward and economically feasible with such coarse station spacing. □

Methods

During the NOAA/OACES I8NR cruise in 1995, stations were occupied along 80° E from 35° S to 5° N with a station spacing of 0.5 to 1 degree. At each station, 24 samples were analysed for a suite of chemical parameters, including DIC, p_{CO_2} , O₂ and chlorofluorocarbons. About half the samples were obtained from the upper 1,000 m, where the strongest anthropogenic CO₂ signal is expected to be present, with the remainder spaced unevenly down to the bottom. DIC samples were analysed using two coulometers equipped with a single operator multi-parameter metabolic analyser¹⁵. The precision of DIC analysis based on replicate analyses of reference material throughout the cruise was $\pm 1.4 \mu\text{mol kg}^{-1}$. Oxygen was analysed by using a modification of the Winkler method¹⁶ and a titrator with a photodiode endpoint detector¹⁷. Dissolved chlorofluorocarbons were measured by electron-capture gas chromatography¹⁸.

During the GEOSECS survey in 1978, DIC and Alk were both measured by means of the potentiometric acid titration method on single samples^{19,20}. The titrators were calibrated using gravimetrically prepared Na₂CO₃ solutions for Alk and DIC²¹. The overall precision of measurement has been estimated as $\pm 8 \mu\text{mol kg}^{-1}$ for Alk and $\pm 11 \mu\text{mol kg}^{-1}$ for DIC²², but these values are subject to systematic errors of unknown origin. This conclusion is based on the internal consistency with the measured Alk and p_{CO_2} data, as well as the consistency with the Atlantic and Pacific GEOSECS data²². The DIC values presented here are taken from GEOSECS Indian Ocean Expedition hydrographic

data 1977–1978 (ref. 23). A re-evaluation of systematic errors between GEOSECS and I8NR data is made by comparing deep-water values to determine the appropriate corrections before making the direct comparison for detecting the anthropogenic carbon signal.

To determine the mean systematic difference between GEOSECS and I8NR, we compared samples from deeper than 2,000 m. First, the I8NR stations are organized into 6 groups corresponding to 6 stations occupied during the GEOSECS. Each group covers a range of 5° latitudes, with the centre location representing the re-occupation of GEOSECS stations. The mean concentration is computed for samples taken at 8 isopycnal surfaces and σ_θ ranges from 27.75 ± 0.01 to 27.82 ± 0.01 , which cover all samples from deeper than 2,000 m. The concentration difference along the same isopycnal surface at the same location between the two cruises is then computed. Results of the comparison are given in Table 1.

We corrected the natural variations of DIC before making the intercomparison. The total carbon released into the water by the respiration of organic matter is estimated using the Redfield ratio¹² C:O₂ of $117 \pm 14:170 \pm 10$. The AOU-corrected DIC (DIC_{AOU}) can be given as $\text{DIC}_{\text{AOU}} = \text{DIC}_o - 0.69 \times \text{AOU}$, where DIC_o is the observed DIC on an isopycnal surface. The choice of this Redfield ratio is based on newer published results and has an insignificant effect on the magnitude of the detected CO₂ signal. If a ratio of 140/172, as derived from Indian Ocean²⁴, is used, the signal will change by $\sim 1 \mu\text{mol kg}^{-1}$ because the anthropogenic DIC increase is derived by difference of two data sets in which the same ratio is applied.

The effect of changes in carbonate dissolution at two different times can be corrected by using Alk data. The DIC corrected for Alk (DIC_{alk}) is given by $\text{DIC}_{\text{alk}} = \text{DIC}_{\text{AOU}} - 0.5 \times (\text{Alk} - \text{Alk}_0)$, where Alk is the observed alkalinity for the isopycnal surface, and Alk₀ is the preformed salinity-normalized Alk, which can be calculated by using the empirical equation²⁵: $\text{Alk}_0 = 2.291 - 2.52(t - 20) + 0.06(t - 20)^2$, where t is the potential temperature at the isopycnal horizon.

To eliminate changes caused by the variations of salt, the corrected DIC is normalized to a salinity of 35.0: $\text{DIC}_n = (\text{DIC}_{\text{alk}}/S) \times 35.0$, where DIC_n is the salinity-normalized final DIC value used for comparison between two different cruises.

For estimating DIC increase from atmospheric CO₂ increase, the following needs to be considered. The potential temperature for waters at isopycnal surfaces between 26.6 and 27.2 ranges from 8 to 12 °C in the equatorial region and 6 to 13 °C in the temperature zone. Taking a mean potential temperature of 10 °C for this upper thermocline water, and a mean surface water alkalinity of $2,290 \mu\text{mol kg}^{-1}$, the DIC at chemical equilibrium with the atmospheric CO₂ can be calculated. The rate of atmospheric CO₂ increase varies from 0.8 p.p.m. yr⁻¹ in the early 1960s to ~ 1.8 p.p.m. yr⁻¹ in recent years^{26,27}. The DIC increase is computed using the rate of atmospheric CO₂ increase of 1.0, 1.5 and 2.0 p.p.m. yr⁻¹.

Received 9 January; accepted 15 September 1998.

- Keeling, C. D. & Whorf, T. in *Trends '93: A Compendium of Data on Global Change* (eds Boden, T., Kaiser, D., Sepanski, R. & Stoss, F.) 16–24 (CDIAC, ORNL, Oak Ridge, Tennessee, 1994).
- Siegenthaler, U. & Sarmiento, J. L. Atmospheric carbon dioxide and the ocean. *Nature* **365**, 119–125 (1993).
- Orr, J. C. *et al.* Transition time for ocean carbon-cycle model comparison. *Research GAIM* **1**, 8–10 (1997).
- Brewer, P. G. Direct observation of the oceanic CO₂ increase. *Geophys. Res. Lett.* **5**, 997–1000 (1979).
- Chen, C.-T. A. & Millero, F. J. Gradual increase of oceanic CO₂. *Nature* **277**, 205–206 (1979).
- Gruber, N., Sarmiento, J. L. & Stocker, T. F. An improved method for detecting anthropogenic CO₂ in the oceans. *Glob. Biogeochem. Cycle* **10**, 809–837 (1996).
- Gruber, N. Anthropogenic CO₂ in the Atlantic Ocean. *Glob. Biogeochem. Cycle* **12**, 165–191 (1998).
- Wallace, D. W. R. Monitoring global ocean carbon inventories. *OOSDP Background Rep. 5* (Ocean Obs. Syst. Dev. Panel, Texas A & M Univ., College Station, 1995).
- Slansky, C. M., Feely, R. A. & Wanninkhof, R. in *Biogeochemical Processes in the North Pacific* (ed Tsunogai, S.) 70–79 (Japan Marine Science Foundation, Tokyo, 1997).
- Bullister, J. L. Chlorofluorocarbons as time dependent tracers in the Ocean. *Ocean. Mag.* **2**, 12–17 (1989).
- Broecker, W. S., Sutherland, S., Smethie, W., Peng, T.-H. & Ostlund, G. Oceanic radiocarbon: separation of the natural and bomb components. *Glob. Biogeochem. Cycle* **9**, 263–288 (1996).
- Anderson, L. A. & Sarmiento, J. L. Redfield ratios of remineralization determined by nutrient data analysis. *Glob. Biogeochem. Cycle* **8**, 65–80 (1995).
- Sarmiento, J. L., Murnane, R. & Le Quére, C. Air–sea CO₂ transfer and the carbon budget of the North Atlantic. *Phil Trans. R. Soc. Lond. B* **348**, 211–218 (1995).
- Murnane, R., Sarmiento, J. L. & Le Quére, C. The spatial distribution of air–sea CO₂ fluxes and the interhemispheric transport of carbon by the oceans. *Glob. Biogeochem. Cycle* (in the press).
- Johnson, K. M., Wills, K. D., Butler, D. B., Johnson, W. K. & Wong, C. S. Coulometric total carbon dioxide analysis for marine studies: maximizing the performance of an automated continuous gas extraction system and coulometric detector. *Mar. Chem.* **44**, 167–189 (1993).
- Carpenter, J. H. The Chesapeake Bay Institute technique for the Winkler dissolved oxygen method.

Limnol. Oceanogr. **10**, 141–143 (1965).

- Friederich, G. E., Sherman, P. & Codispoti, L. A. A high precision automated Winkler titration system based on a HP-85 computer, a simple colorimeter and an inexpensive electromechanical buret. *Bigelow Lab. Tech. Rep.* **42**, 1–24 (1984).
- Bullister, J. L. & Weiss, R. F. Determination of CCl₃F and CCl₂F₂ in seawater and air. *Deep-Sea Res.* **35**, 839–853 (1988).
- Bradshaw, A. L., Brewer, P. G., Shafer, D. K. & Williams, R. T. Measurements of total carbon dioxide and alkalinity by potentiometric acid titration in the GEOSECS program. *Earth Planet. Sci. Lett.* **55**, 99–115 (1981).
- Brewer, P. G., Bradshaw, A. L. & Williams, R. T. in *The Changing Carbon Cycle: A Global Analysis* (eds Trabalka, J. R. & Reichle, D. E.) 348–370 (Springer, New York, 1986).
- Takahashi, T. & Bainbridge, A. E. in *GEOSECS Atlantic Expedition Vol. 1* (ed. Bainbridge, A. E.) 7–10 (US Gov. Printing Office, Washington DC, 1981).
- Takahashi, T. in *GEOSECS Indian Ocean Expedition Vol. 5* (eds Weiss, R., Broecker, W. S., Craig, H. & Spencer, D.) 5–7 (US Gov. Printing Office, Washington DC, 1983).
- Weiss, R., Broecker, W. S., Craig, H. & Spencer, D. *GEOSECS Indian Ocean Expedition Vol. 5* (US Gov. Printing Office, Washington DC, 1983).
- Takahashi, T., Broecker, W. S. & Langer, S. Redfield ratio based on chemical data from isopycnal surfaces. *J. Geophys. Res.* **90**, 6907–6924 (1985).
- Millero, F. J., Lee, K. & Roche, M. Distribution of alkalinity in the surface waters of the major oceans. *Mar. Chem.* **60**, 111–130 (1988).
- Conway, T. J. *et al.* Atmospheric carbon dioxide measurements in the remote global troposphere 1981–1984. *Tellus* **40B**, 81–115 (1988).
- Conway, T. J. *et al.* Evidence for interannual variability of the carbon cycle from the NOAA/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network. *J. Geophys. Res.* **99**, 22831–22855 (1994).

Acknowledgements. We thank the officers and crew of the NOAA ship *Malcolm Baldrige* for their assistance; T. Lantry, M. Roberts, H. Chen and D. Greeley for DIC and f_{CO_2} analyses and data reduction; D. Wisegarver, F. Menzies and D. Greeley, for CFC analysis; G. Thomas, D. Anderson and R. Roddy for CTD operation, and oxygen and salinity analyses; R. Molinari and A. Pfaff for CTD data reduction; T. Hughes for providing the DIC results from Princeton annual mean ocean GCM; and J. Orr for constructive comments. This research was funded by the NOAA Climate and Global Change (C&GC) program and the NOAA Environmental Research Laboratories. We thank J. Todd of the C&GC office for supporting the NOAA component of the WOCE repeat hydrography.

Correspondence and requests for materials should be addressed to T.-H.P.

Passive infrared spectroscopy of the eruption plume at Popocatepetl volcano, Mexico

S. P. Love*, F. Goff*, D. Counce*, C. Siebe† & H. Delgado†

* NIS and EES Divisions, Los Alamos National Laboratory, Mail Stop C323, Los Alamos, New Mexico 87545, USA

† Instituto de Geofísica, Universidad Nacional Autónoma de México, Coyoacán 04510, México, DF, México

Volcanic gases provide important insights into deep-Earth processes, and gas composition and flux variations show promise as predictors of eruptive activity^{1–3}. But data correlating gas composition with eruptions are sparse, largely because such studies have traditionally involved direct sampling inside a volcanic crater—a hazardous operation that has resulted in numerous deaths^{4,5}. Crater-rim-based spectroscopy^{6–9}, closed-path spectroscopy of gases sampled from aircraft¹⁰, and time-averaged studies using volatile traps^{11–13} allow measurements to be taken from safer distances. But when a full-scale explosive eruption threatens, even these methods become dangerous as the hazard radius expands to many kilometres. Previously, only sulphur dioxide has been reliably measurable at such large distances, using correlation spectroscopy¹⁴. Here we describe techniques that extend the useful range of passive infrared spectroscopy to monitor many gases at distances of over 17 km. We demonstrate the use of these techniques in a high-temporal-resolution study of short-term compositional variations associated with an explosive eruption at Mexico's Popocatepetl volcano on 25–26 February 1997. We observed a steady increase in SiF₄/SO₂ over several days preceding the eruption, followed by a tenfold decrease in this ratio over a few hours immediately afterwards.

Open-path infrared spectroscopy of volcanic gases (as opposed to closed-path techniques requiring gas to be drawn into the instrument) has been demonstrated at a number of volcanoes by several groups. But this technique has always required rather special circumstances—either fortuitously located high-temperature rocks

or lava^{6–9,15–18} acting as an infrared source, or a combination of accessible topography and minimal volcanic hazard allowing installation of an artificial infrared source⁶ across the plume from the spectrometer. Such optimal conditions are not, however, generally present at the most active and dangerously explosive volcanoes for which gas monitoring is most crucial. Here we demonstrate techniques that utilize the strong radiance contrast between the gas plume and the sky to obtain high-quality infrared spectra in a truly passive manner, without artificial infrared sources, from distances of many kilometres. Unlike hot-source measurement strategies, these sky-background techniques are applicable to

almost any volcano, and permit infrared spectra to be obtained from much safer distances.

Popocatepetl (5,452 m), 70 km southeast of Mexico City, is a recently reawakened calc-alkaline andesite/dacite; stratovolcano^{19,20}, and provides an ideal target for demonstrating the utility of these techniques. Over the past 23,000 years, it has erupted explosively in cycles of 1,000–3,000 years, with large eruptions between 3195 and 2830 BC, 800 and 215 BC, and AD 675 and 1095 (ref. 21). Activity in the 1920s emplaced an intracrater dacitic dome 215 m high²². The current explosive phreato-magmatic activity started in December 1994 (refs 23, 24), temporarily displacing 75,000 people. A new

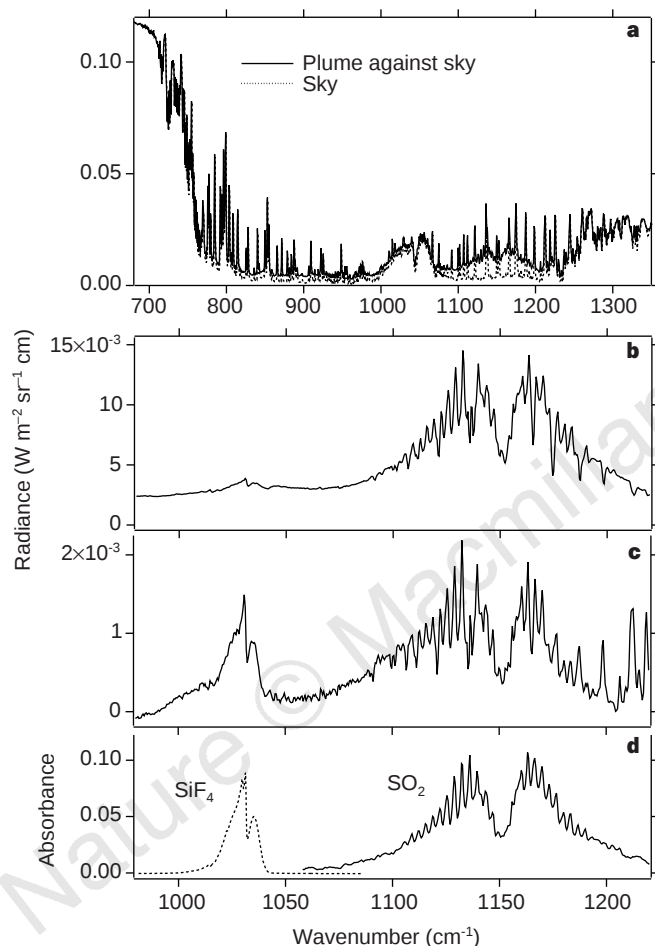


Figure 1 Radiometrically calibrated clear-sky passive LWIR spectra of the Popocatepetl plume, 1 cm⁻¹ resolution. All LWIR spectra were obtained using a liquid-N₂-cooled HgCdTe detector. **a**, In a typical data collection, a spectrum of the gas plume viewed against the sky is first obtained (solid line), followed by a spectrum of clean sky upwind of the plume, at the same elevation angle (dotted line). Shown here are spectra obtained on 23 February 1997 from the Tlamacaz site, 4 km from the summit crater. Prominent in both spectra is an emission feature extending from roughly 1,000 to 1,070 cm⁻¹ arising from stratospheric ozone, while the most obvious divergence between the two curves occurs at the plume's SO₂ emission feature, extending from roughly 1075 to 1225 cm⁻¹. **b**, Difference between the spectra in **a**. For optically thin plumes, this difference is (to first order) the radiance spectrum of the plume gases alone. We note the weak SiF₄ feature at 1,030 cm⁻¹. **c**, Another difference, of a spectrum pair obtained on 26 February 1997 from 10 km from the summit, showing a prominent SiF₄ feature. The ratio of SiF₄ to SO₂ varies dramatically from day to day as the eruptive state of the volcano changes (see Fig. 3); this spectrum corresponds to the maximum value of SiF₄/SO₂ observed during this campaign. **d**, Laboratory absorption spectra³⁰ of SiF₄ and SO₂, shown for comparison.

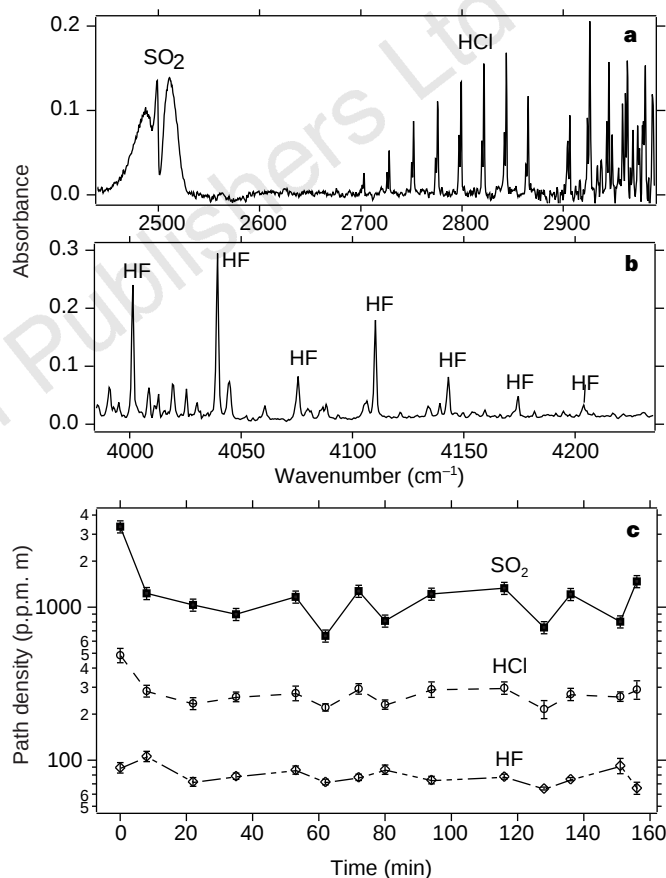


Figure 2 Mid-wave infrared results for the Popocatepetl plume. **a**, Absorption spectrum obtained on 27 February 1997 from the Tlamacaz site using sunlight forward-scattered by a high cirrus overcast, resolution 1 cm⁻¹. SO₂ and HCl features are noted. The cloud-scattered sunlight technique offers substantial signal-to-noise improvement over simple thermal emission: at a wavelength of 4 μm, (2,500 cm⁻¹), cloud-scattered sunlight is as much as 20 times brighter than thermal emission from the plume, and at 2.5 μm it is several orders of magnitude brighter. Clearly seen here are the 2,500 cm⁻¹ SO₂ vibrational combination band, and the series of nearly equally spaced lines of the HCl vibrational-rotational spectrum, which extends from roughly 2,700 cm⁻¹ to over 3000 cm⁻¹, where it eventually is obscured by atmospheric absorption. The two components of each HCl line, just discernible in this figure, arise from ³⁵Cl and ³⁷Cl isotopes. **b**, Continuation of the spectrum of **a** into the near-infrared region, showing HF absorption features. **c**, Time dependence of the path densities of SO₂, HCl and HF, derived from FTIR absorption spectra obtained from the Tlamacaz site during a 2.5-hour period beginning at 12:30 pm local time on 27 February 1997. Values are based on absorption spectra, as in **a**, **b**, obtained using the cloud-scattered sunlight technique. As data points were obtained by repeatedly performing measurements with the instrument aimed at the same location over the summit crater, they represent variations in plume density at a particular point in space, rather than variations in integrated flux. All MWIR spectra were obtained using a liquid-N₂-cooled InSb detector.

intracrater hornblende dacite dome, 62–64 wt% SiO₂ (ref. 25), grew from March 1996 to July 1997. Frequent explosions have continued to date, producing ash plumes as high as 10 km above the summit²⁶, precluding any gas-measurement technique requiring human presence near the summit.

In February 1997, we performed remote infrared measurements at Popocatepetl using a commercial (Midac, Irvine, CA) Fourier-transform infrared (FTIR) spectrometer, slightly modified to improve passive performance. We obtained long-wave infrared (LWIR, 8–12 µm) spectra using only thermal emission from the gases themselves, viewed against the colder backdrop of the sky, a strategy equally suited to night or day measurements. Clear sky provides the best thermal contrast, but overcast conditions produce acceptable results. Figure 1 displays typical LWIR clear-sky spectra, showing SO₂, numerous H₂O lines, plume aerosol continuum emission, and (notably) SiF₄. SiF₄ was detected at Italy's Vulcano and Etna volcanoes using FTIR with a hot infrared source (volcanically heated rock or an artificial source) by Francis *et al.*⁶, who pointed out the potential importance of this gas for inferring fumarole temperatures. Now, using the sky-background method, we can monitor SiF₄ from much greater distances. Absolute radiometric calibration, however, is essential for quantitative measurements with this technique. This was performed approximately every hour using two portable blackbody sources consisting of insulated, temperature-controlled, grooved aluminium plates coated with a high-infrared-emissivity paint²⁷. A 30-cm aperture-filling design permits determination and removal of the substantial instrument self-emission.

Mid-wave infrared (MWIR, 3–5 µm) spectra can be similarly obtained (we have successfully measured volcanic HCl this way²⁸), but much greater sensitivity is obtained by viewing a high, thin cloud layer in the general direction of the Sun (within ~40°). Such clouds strongly scatter sunlight towards the instrument, providing an ideal bright background against which the plume can be viewed in absorption. In Fig. 2, absorption spectra obtained using cloud-scattered sunlight clearly show SO₂, HCl and HF. This is, to our knowledge, the first remotely measured volcanic HF spectrum (an HF/SiO₂ value based on this spectrum first appeared in a report by Goff *et al.*¹³).

Supplementing the FTIR measurements, we measured the total SO₂ flux with the widely used ultraviolet correlation spectrometer, or Cospec (see Kyle *et al.*²⁹ for the method). FTIR measures relative amounts of many gases, but requires knowledge of gas temperature to determine absolute amounts; in contrast, the Cospec, using scattering sunlight, measures SO₂ only, but in a temperature-independent and absolute fashion. To utilize these complementary characteristics, FTIR and Cospec were deployed side-by-side at two sites for the data shown here: (1) Tlmacaz lodge (elevation 3,900 m), a map distance of 4 km from the summit, and (2) west

of the village of Santiago Xalitintla, 10 km from the summit, at elevation 2,900 m.

FTIR-derived ratios³⁰ of the various gases are summarized in Table 1, together with earlier (1994–96) crater-rim volatile trap results obtained using KOH traps that absorbed acid volatiles and volatile metals¹³. Trap-based HCl/SO₂ and HF/SO₂ ratios are similar to the FTIR results, providing independent verification. Popocatepetl's SO₂ flux varies dramatically, but averaged about 2,000 tonnes per day (t d⁻¹) in 1994, 1,600 t d⁻¹ in 1995, and 15,000 t d⁻¹ in 1996 (ref. 13), occasionally exceeding 50,000 t d⁻¹ in 1997 (1 t d⁻¹ = 0.01157 kg s⁻¹). Fluxes of other acid gases are surprisingly large, thousands and hundreds of tonnes per day of HCl and HF, respectively. SiF₄ flux is typically a few tonnes per day, but can exceed 1,000 t d⁻¹ during more intense activity. Popocatepetl's average SiF₄/SO₂ value of 2.9×10^{-3} lies between the values of 3.7×10^{-4} and 5.9×10^{-3} reported for Etna (basalt-hawaiite) and Vulcano (rhyolite), respectively⁶, consistent with the expected higher F concentration in more silicic melts.

Variations in SO₂, HCl and HF over a 2.5-hour period are plotted in Fig. 2c, illustrating the ability of infrared remote sensing to monitor short-term composition fluctuations. We note that the relative amplitude of the concentration fluctuations is not the same for all gases, perhaps reflecting the actual fumarolic output, or possibly resulting from processes occurring within the plume (for example, dissolution of the highly soluble HCl and HF within liquid water droplets in the plume, or segregation by weight).

Perhaps the most intriguing data obtained during this campaign was a series of 43 LWIR spectra obtained over four days (21–24 and 26 February 1997) culminating in ash eruptions on 25–26 February. The steady increase in SiF₄/SO₂ preceding the eruption, its extremely high value immediately afterwards, and subsequent rapid decline are summarized in Fig. 3, together with the total SO₂ flux.

Previous research demonstrates that SiF₄ increases with falling temperature and is not abundant at magmatic temperatures^{31,32}. The effective equilibrium temperature of the volcanic gases can be calculated from the SiF₄/HF ratio assuming the following reaction^{6,31}: SiO₂ + 4HF → SiF₄ + 2H₂O, where SiO₂ is magmatic silica glass. Using data from Symonds *et al.*³³ (their equation (7) and constants from their Appendix 2), our FTIR-measured gas ratios, and assuming a molar CO₂/SO₂ ratio of 2.75 and H₂O mole fraction of 0.9 (ref. 13), we calculate a temperature of 200 ± 20 °C for the observed range of gas ratios. This is clearly sub-magmatic and must indicate reaction of the gases with the conduit walls during passage to the surface, or, during explosions, reaction with ash particles in the rapidly cooling plume.

Figure 3 is somewhat surprising in that it implies gas cooling before the explosion, with a minimum temperature immediately afterwards. We speculate that this reflects an adiabatic expansion which occurred as a plug in the volcanic conduit began giving way

Table 1 Chemical ratios and flux rates of Popocatepetl plume components 1994–97

Date	HCl/SO ₂ weight ratios		HF/SO ₂ weight ratios		SiF ₄ /SO ₂ weight ratios FTIR	Cospec (tons d ⁻¹) SO ₂	Approximate flux rate (t d ⁻¹)			
	Trap	FTIR	Trap	FTIR			HCl	HF	SiF ₄	
Jul. 1994	0.07†	–	0.05†	–	–	3,100†	230†	170†	–	
Dec. 1994	0.098†	–	0.025†	–	–	3,960†	400†	125†	–	
Mar. 1996	0.21†	–	0.021†	–	–	12,900†	2,700†	270†	–	
21 Feb. 1997	–	–	–	–	0.0017	4,100*	533†	78†	7.0‡	
23 Feb. 1997	–	0.11	–	0.015	0.0027	2,700*	297†	41†	7.3‡	
24 Feb. 1997	–	–	–	–	0.0039	2,000*	260†	38†	4.6‡	
26 Feb. 1997	–	–	–	–	0.018–0.0017	60,000†–4,000	7,800†–520†	1,140†–76‡	1,080†–7‡	
27 Feb. 1997	–	0.13†	–	0.019†	0.0	13,000*†	1,700†	250†	–	

* Average value measured on the day given.

† From Goff *et al.*¹³.

‡ Values calculated by multiplying measured ratio by the SO₂ flux rate. For days on which HCl and HF were not measured, values are calculated using the 2/27 ratios. Quantitative gas ratios were obtained by comparing the observed spectra with a database of published absorption spectra³⁰ at the same spectral resolution. For LWIR radiance spectra, database spectra are first converted from absorbance to emissivity, then to spectral radiance by multiplying by the Planck function evaluated at the summit air temperature, then fitted to the observed radiance by varying the gas amount. Note that since SO₂ and SiF₄ features lie at nearby wavenumbers, near the peak of the Planck spectrum, the SiF₄/SO₂ ratio is insensitive to plume temperature errors; an error of 20 K results in only a 3.8% error in the SiF₄/SO₂ ratio. A simple absorption comparison was used for the MWIR scattered-sunlight absorption spectra (HCl/SO₂ and HF/SO₂).

–, Not available.

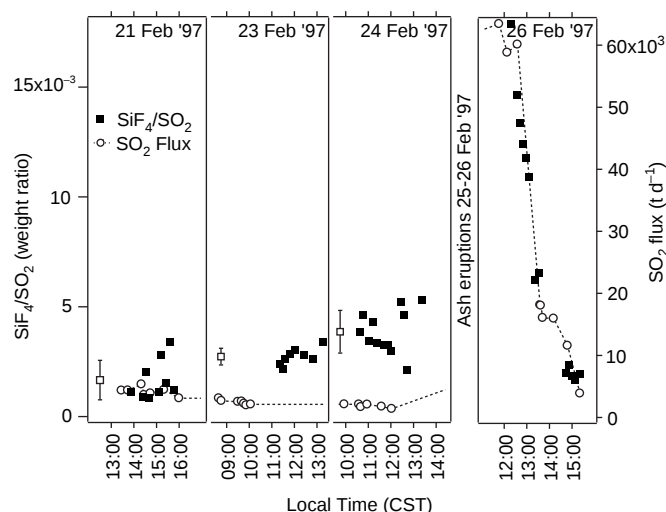


Figure 3 Time variation of SiF_4 and SO_2 output at Popocatepetl, 21–26 February 1997. Shown here are the FTIR-derived SiF_4/SO_2 ratio (filled squares) and Cospec-derived total SO_2 flux (open circles, dashed line) as a function of time over several days preceding and immediately after a series of moderate explosive ash eruptions. On the days preceding the eruptions, the molar SiF_4/SO_2 ratio increased steadily, from 21 February 1997 when the mean SiF_4/SO_2 ratio was 1.7×10^{-3} (standard deviation $\sigma = 9 \times 10^{-4}$) through to 23 February (mean, 2.7×10^{-3} ; $\sigma = 4 \times 10^{-4}$), to 24 February (mean, 3.9×10^{-3} ; $\sigma = 9.7 \times 10^{-4}$). Open squares show the average SiF_4/SO_2 ratio for the days preceding the eruption; error bars indicate standard deviation of the intraday fluctuations, $\pm\sigma$. From 21 to 24 February, the daily average SiF_4/SO_2 undergoes a statistically significant increase of more than 2σ (note that the bars for these two days do not overlap). During the same period the total SO_2 flux decreased somewhat, from roughly $4,000$ to $2,000 \text{ t d}^{-1}$. On 26 February, immediately after the explosive ash event, SiF_4/SO_2 stood at 1.1×10^{-2} and the SO_2 flux was $>60,000 \text{ t d}^{-1}$. Over the course of 3 hours, both values dropped precipitously back to their pre-eruption levels of 21 February.

before the ash-laden explosion; as the plug weakened, releasing gas, pressure and temperature decreased, increasing SiF_4/SO_2 . Eventually the plug failed catastrophically, producing the observed explosion, a large drop in temperature and pressure, and the observed spike in SiF_4/SO_2 . A new constriction then developed, returning SiF_4/SO_2 to earlier levels. Such an explosion, driven by gas build-up rather than significant magma movements, is consistent with the overall low gas temperatures implied by the SiF_4/SO_2 ratios and the relatively minor nature of this eruption. Future observations of similar SiF_4/SO_2 trends could warn scientists of similar explosions, which occur frequently at Popocatepetl.

Passive spectroscopy can be extended to ranges considerably beyond the 10 km used here. We recently obtained spectra from 17.5 km, with high data quality implying that even greater ranges are possible. Weather ultimately limits the useful range; radiation at these wavelengths cannot penetrate clouds, and absorption by atmospheric water vapour reduces the sensitivity with increasing distance. Dense ash during large eruptions also severely limits sensitivity. The technique works best at high-altitude, dry locations, where ranges of 30 km or more should be achievable. In wetter locations, ranges will be shorter and clear days suitable for measurements less frequent. But compared to the ultraviolet, infrared is less sensitive to particulate pollution as scattering decreases with increasing wavelength; we obtained good FTIR data even when smoke from local brush fires rendered the Cospec unusable.

These infrared techniques complement the existing suite of technologies available for gas monitoring, and are uniquely valuable at explosive volcanoes like Popocatepetl. As we have demonstrated even in this limited study, time series of infrared spectra permit significant insights into eruptive processes. With a range greater

than other multi-gas techniques, and unsurpassed temporal resolution, these infrared techniques permit quasi-continuous gas composition monitoring right up to an eruption, without unduly endangering the scientists involved. The eruption of Popocatepetl on 25–26 February 1997 would have killed scientists attempting to obtain data like that of Fig. 3 by either direct sampling or crater-rim based methods, and would have destroyed any *in situ* telemetry-based instruments, precluding post-eruption observations. Even aircraft-based measurements would have entailed substantial risk. In less hazardous locations, volatile traps and airborne sampling can provide information that is difficult (CO_2) or impossible (for example, metals or isotopes) for remote FTIR. Direct sampling permits the most exhaustive analysis, and is practical at the lowest-activity levels, when gas emission rates become inconveniently low for remote sensing. But for large emitters, remote FTIR safely provides a superior spatial average of the total plume composition, which, when combined with Cospec, allows total flux determination for many important volcanic gases. Future work should emphasize longer-term, more-continuous gas monitoring, ideally combining remote FTIR with other techniques as safety permits, along with seismic and geodesic measurements, in order to correlate more fully trends in gas composition with various types of eruptive activity. □

Received 18 December 1997; accepted 25 September 1998.

- Allard, P. in *Forecasting Volcanic Events* (eds Tazieff, H. & Sabroux, J.-C.) 337–394 (Elsevier, Amsterdam, 1983).
- Fischer, T. P. et al. Correlation between SO_2 flux and long-period seismicity at Galeras volcano. *Nature* **368**, 135–137 (1994).
- Tedesco, D. *Monitoring Active Volcanoes* (eds McGuire, B., Kilburn, C. & Murray, J.) (UCL Press, London, 1995).
- Kerr, R. A. Volcanologists ponder a spate of deaths in the line of duty. *Science* **260**, 289–290 (1993).
- Baxter, P. J. & Gresham, A. Deaths and injuries in the eruption of Galeras Volcano, Colombia, 14 January 1993. *J. Volcanol. Geotherm. Res.* **77**, 325–338 (1997).
- Francis, P., Chaffin, C., Maciejewski, A. & Oppenheimer, C. Remote determination of SiF_4 in volcanic plumes: A new tool for volcano monitoring. *Geophys. Res. Lett.* **23**, 249–252 (1996).
- Francis, P., Maciejewski, A., Oppenheimer, C., Chaffin, C. & Caltabiano, T. SO_2/HCl ratios in the plumes from Mt. Etna and Vulcano determined by Fourier transform spectroscopy. *Geophys. Res. Lett.* **22**, 1717–1720 (1995).
- Mori, T. & Notsu, K. Remote CO , COS , CO_2 , SO_2 , HCl detection and temperature estimation of volcanic gas. *Geophys. Res. Lett.* **24**, 2047–2050 (1997).
- Stix, J., Morrow, W., Nicholls, R. & Charland, A. Infrared remote sensing of CO and COS gas emitted by the Galeras Volcano, Colombia, January 8–10, 1993. *Can. J. Remote Sens.* **22**, 297–304 (1996).
- McGee, K. A. & Gerlach, T. M. Airborne volcanic plume measurements using a FTIR spectrometer, Kilauea volcano, Hawaii. *Geophys. Res. Lett.* **25**, 615–618 (1998).
- Noguchi, K. & Kamiya, H. Prediction of eruptions by measuring the chemical composition and amount of gases. *Bull. Volcanol.* **24**, 367–378 (1963).
- Italiano, F., Nuccio, P. M., Pecoraino, G. & Valenza, M. Monitoring of acid gases at the crater of Vulcano using the method of the alkaline traps. *Acta Vulcanol.* **1**, 249–254 (1991).
- Goff, F. et al. Geochemical surveillance of magmatic volatiles at Popocatepetl Volcano, Mexico. *Geol. Soc. Am. Bull.* **110**, 695–710 (1998).
- Realmuti, V. J., Sutton, A. J. & Elias, T. Multispectral thermal infrared mapping of sulfur dioxide plumes: A case study from the East Rift Zone of Kilauea Volcano, Hawaii. *J. Geophys. Res.* **B 107**, 15057–15072 (1997).
- Naughton, J. J., Derby, J. V. & Glover, R. B. Infrared measurements on volcanic gas and fume: Kilauea eruption, 1968. *J. Geophys. Res.* **74**, 3273–3277 (1969).
- Mori, T., Notsu, K., Tohjima, Y. & Wakita, H. Remote detection of HCl and SO_2 in volcanic gas from Unzen volcano, Japan. *Geophys. Res. Lett.* **20**, 1355–1358 (1993).
- Mori, T. et al. Remote detection of fumarolic gas chemistry at Vulcano, Italy, using an FT-IR spectral radiometer. *Earth Planet. Sci. Lett.* **134**, 219–224 (1995).
- Notsu, K., Mori, T., Igarashi, G., Tohjima, Y. & Wakita, H. Infrared spectral radiometer: A new tool for remote measurement of SO_2 of volcanic gas. *Geochem. J.* **27**, 361–366 (1993).
- Nixon, G. T., Demant, A., Armstrong, R. L. & Harakal, J. E. K-Ar and geologic data bearing on the age and evolution of the Trans-Mexican Volcanic Belt. *Geofis. Int.* **26**, 109–158 (1987).
- Robin, C. Le volcan Popocatepetl (Mexique): structure, evolution petrologique et risques. *Bull. Volcanol.* **47**, 299–315 (1984).
- Siebe, C., Abrams, M., Macias, J. L. & Obenholzer, J. Repeated volcanic disasters in Prehispanic time at Popocatepetl, central Mexico: Past key to the future? *Geology* **24**, 399–402 (1996).
- Waiz, P. Popocatepetl again in activity. *Am. J. Sci.*, 5th ser. **1(1)**, 81–83 (1921).
- Delgado, H. & Cárdenas, G. L. in *Volcanic Activity and the Environment* (abstr.) 49 (IAVCEI, Puerto Vallarta, Mexico, 1997).
- Delgado, H., Venegas, J. J., Cárdenas, L. & Doukas, M. Explosive eruption hazards at Popocatepetl Volcano (Mexico): The SO_2 flux measurements insight to the problem. *Geol. Soc. Am. Annu. Meeting, Prog. Abstr.* A-109 (1995).
- Stimac, J. A., Siebe, C. & Schaaf, P. in *Volcanic Activity and the Environment* (abstr.) 76 (IAVCEI, Puerto Vallarta, Mexico, 1997).
- De la Cruz, S. & Siebe, C. The giant Popocatepetl stirs. *Nature* **388**, 227 (1997).
- Fowler, J. B. A third generation water bath based blackbody source. *J. Res. Natl. Inst. Stand. Technol.* **100**, 591–599 (1995).
- Love, S. P. et al. Integrated infrared/UV/LIDAR remote sensing of volcanic emissions at White island and Mt. Ruapehu, New Zealand. (Abstr.) *Eos* **77**, F802 (1996).
- Kyle, P. R., Sybelton, L. M., McIntosh, W. C., Meeker, K. & Symonds, R. in *Volcanological and Environmental Studies of Mount Erebus, Antarctica* (ed. Kyle, P. R.) 69–82 (Am. Geophys. Union, Washington DC, 1994).

30. Hanst, P. L., Hanst, S. T. & Williams, G. M. *Infrared Spectra for Quantitative Analysis of Gases* (QASoft Database) (Infrared Analysis, Inc., Anaheim, CA, 1996).
31. White, A. F. & Hochella, M. F. Surface chemistry associated with the cooling and subaerial weathering of recent basalt flows. *Geochim. Cosmochim. Acta* **56**, 3711–3721 (1992).
32. Symonds, R. B., Reed, M. H. & Rose, W. I. Origin, speciation, and fluxes of trace-element gases at Augustine volcano, Alaska: insights into magma degassing and fumarolic processes. *Geochim. Cosmochim. Acta* **56**, 633–657 (1992).
33. Symonds, R. B. & Reed, M. H. Calculation of multicomponent chemical equilibria in gas-solid-liquid systems: Calculation methods, thermochemical data and applications to studies of high-temperature volcanic gases with examples from Mount St. Helens. *Am. J. Sci.* **293**, 758–864 (1993).

Acknowledgements. We thank R. Meli for access and logistical support, and P. G. Weber, S. Gerstl and D. Pettit for discussions on remote sensing applications. This work was supported by Laboratory Directed Research and Development (LDRD) grants (Remote Sensing Science) from Los Alamos National Laboratory, and by the Mexican National Center for Disaster Prevention (CENAPRED), the Mexican National Science and Technology Commission (CONACYT), and the Universidad Nacional Autónoma de México.

Correspondence and requests for materials should be addressed to S.P.L. (e-mail: splove@lanl.gov).

Remote measurements of volcanic gas compositions by solar occultation spectroscopy

Peter Francis^{*‡}, Mike R. Burton^{†‡} & Clive Oppenheimer^{†‡}

^{*} Department of Earth Sciences, The Open University, Milton Keynes MK7 6AA, UK

[†] Department of Geography, University of Cambridge, Downing Place, Cambridge CB2 3EN, UK

[‡] All these authors contributed equally to this work

Volcanic gases have important effects on the atmosphere and climate^{1,2} and are important indicators of subsurface magmatic processes^{3,4}, but they are difficult to measure. *In situ* sampling on volcanoes can provide detailed information^{5–7} but is often impractical or hazardous. It is safer to apply remote techniques, for example correlation spectroscopy⁸, which is now widely used to estimate emission rates of sulphur dioxide; but making remote measurements of other gas species has proved more difficult. Developments in Fourier-transform infrared spectroscopy, however, have shown promise^{9–11}. Here we report Fourier-transform infrared observations of volcanic plume compositions that we obtained by solar occultation at Mount Etna in 1997. We found molar ratios of SO₂:HCl and SO₂:HF to be ~4.0 and 10, corresponding to emission rates of HCl and HF of about 8.6 and 2.2 kg s⁻¹, respectively, confirming Mount Etna as the largest known sustained point source of these gases. Solar occultation spectroscopy has advantages over other methods as it enables measurement of plume compositions several kilometres downwind, without requiring hot rocks or lamp sources. The regular and frequent observation of volcanic gases provides a valuable tool for volcano surveillance, and data from plumes at different distances downwind of a volcano's summit may help us to understand the atmospheric chemistry involved in plume dispersal.

A range of geophysical, geochemical and geodetic techniques is used in the routine surveillance of active volcanoes. Of these, seismic monitoring has always provided the foundation. Volcanic gases have been studied both as indicators of eruptive activity, and in order to investigate the role of volatiles in magma evolution⁶. Many different approaches have been used, with technical innovations providing new instruments. Direct sampling of gases followed by laboratory chemical analysis provides the most detailed information^{5–7}, but has obvious limitations. Airborne sampling and sensing provide safer alternatives but are restricted because of the high costs of operating aircraft. Correlation spectrometry (Cospec)⁸, filter packs¹² and infrared analytical instruments¹³, including closed-path Fourier-transform infrared (FTIR) spectrometers¹⁴, have all

been deployed on aircraft. Space-borne sensors, notably the total ozone mapping spectrometer (TOMS), have been used to constrain volcanic SO₂ budgets¹⁵ but are primarily applicable to observations of larger explosive eruptions.

Ground-based remote sensing techniques provide a further approach. Most of the SO₂ emission data for volcanoes around the world have been obtained by making vehicle traverses with Cospec beneath the volcanic plume⁸. Open-path FTIR spectroscopy, extensively used for upper atmosphere research¹⁶, has recently been successfully applied to study gas emissions from several volcanoes. Whereas closed-path FTIR systems¹⁴ use a spectrometer with a folded optical path into which plume gases are pumped, open-path FTIR systems view a natural or artificial infrared source along a path that intersects the gas plume. During earlier experiments with FTIR spectroscopy, we and others have successfully used lamps or hot rocks as infrared sources^{9–11} but field circumstances often dictate that it is impractical to contrive a suitable, safe geometry of spectrometer, gas plume and source.

Direct solar radiation provides an intense source across much of the wavelength range of interest and has been used in many FTIR spectroscopic studies of the stratosphere¹⁶. We investigated the volcanological potential of this methodology on Mount Etna (~3,300 m summit altitude), a prodigious source of gases¹⁷. Etna's pattern of persistent, frequently effusive activity has made it an exceptional natural laboratory for developing volcanological techniques for decades. Since 1975, its SO₂ output has been measured by Cospec and found to average about 55 kg s⁻¹ (1 kg s⁻¹ = 86.4 t d⁻¹); variations in this emission rate have been correlated with levels of eruptive activity^{17–19}.

During June and October 1997, a period characterized by mildly explosive activity in the Bocca Nuova and Southeast crater, with minor lava effusion from the latter, we made solar occultation observations of Etna's plume from Zafferana, 12 km distant from the volcano (Fig. 1a, b) and a site near the Torre dell Filosofo, 1 km from the summit craters (Fig. 1c, d). These sites were chosen to locate the Sun behind the plume, the Sun being viewed directly by the spectrometer and tracked between measurements by manual adjustment of tripod gears. Spectra were collected with both indium antimonide (InSb) and mercury cadmium telluride (MCT) detectors, practical frequency ranges for solar measurements being 2,000–6,000 and 500–5,000 cm⁻¹, respectively. We also made comparative observations using a short lava flow erupted from the Southeast crater as an infrared source (Fig. 1e).

We retrieved gas column amounts (p.p.m. m) from single beam spectra employing a non-linear least squares fitting procedure and a forward model that simulates an open path spectrum. The forward model assumes a single pressure and temperature for the plume gases. Spectral windows at 2,450–2,550, 2,700–2,850 and 4,050–4,150 cm⁻¹ were selected to retrieve SO₂, HCl and HF, respectively. In all retrievals the target gases were fitted simultaneously with a background correction and a frequency shift.

SO₂, HCl and HF column amounts are strongly correlated, and there is good agreement between the two sets of solar data and those obtained using the lava flow source (Fig. 1e). The SO₂:HCl molecular ratios we measured in both June and October 1997 (4.0 and 4.6, respectively; Fig. 1a, c) are in the same range as those that we measured in September 1994 (~3.0–5.0) using a lamp source and an alternative retrieval method¹⁰. Although direct comparisons are not straightforward because of differences in measurement techniques, the equivalent Cl:S and Cl:F mass ratios that we derive (0.24–0.27 and 3.8–6.6, respectively) lie within the ranges reported for filter pack studies by Pennisi and LeCloarec²⁰. They deployed filter packs at the Bocca Nuova, Voragine and Southeast crater vents sporadically between 1992 and 1995, and their results suggested that during periods of persistent deep degassing (~100 MPa), Cl:S mass ratios exceed 1, whereas during eruptive episodes they are around 0.1. They attributed this to strong sulphur exsolution in shallow,

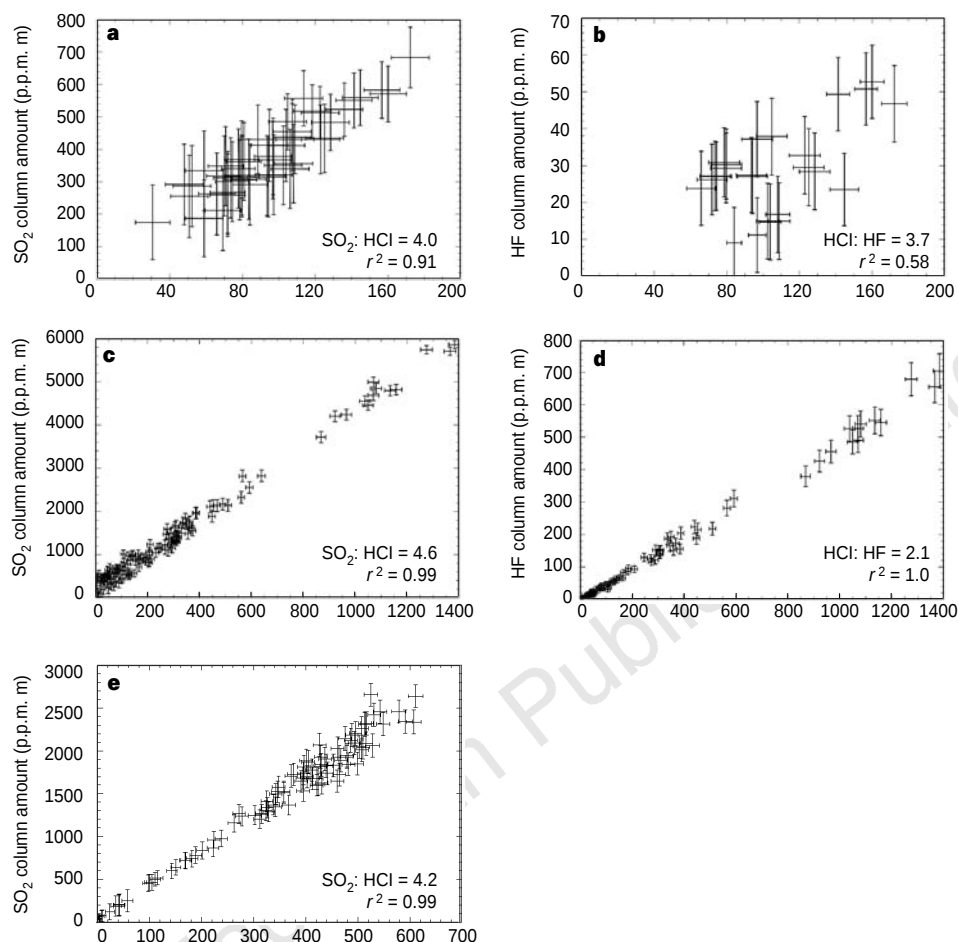


Figure 1 Concentration-path lengths of SO₂, HF and HCl in the Etna plume measured by FTIR spectrometry. **a, b**, Data were measured from Zafferana, 12 km from the volcano using the Sun as the infrared source, 14 June 1997. **c, d**, Data were measured from the Torre dell Filosofo, about 1,000 m from the summit vents, also using the Sun, 26 October 1997. Comparison of **a** and **b** with **c** and **d** shows that measured column amounts were one order of magnitude lower in the dilute

downwind plume than near the vent, making retrievals more difficult. Simultaneous observations would permit investigations of down-plume chemistry. **e**, SO₂ versus HCl column amounts obtained using an incandescent lava flow from the southeast crater as the infrared source, measured from the Torre dell Filosofo, 3 November 1997. In each case, the best-fit linear regression yields the molecular ratio of the two gases represented.

partially degassed magma. Our 1997 measurements of low Cl:S (Table 1) appear consistent with this two-stage degassing model because they correspond to a period of lava effusion from the Southeast crater, which preceded much more vigorous activity in December 1997 and January 1998.

Filter pack studies have also suggested that individual vents exhibit characteristic Cl:F ratios; the Bocca Nuova ratio being a factor of two higher than the Voragine, which has a ratio roughly twice that of the Southeast crater^{18,20}. Because some filter packs determine total S and Cl, whereas FTIR spectroscopy determines gaseous SO₂ and HCl exclusively, caution is necessary in comparing results. Phases such as KCl and NaCl may be abundant in Etna gases²¹, and would not be sensed by our technique. Our Cl:F estimates probably represent contributions from both vents but provide no evidence for mixing of sources with different Cl:F signatures as might be expected. HF retrievals are most robust at near vent sites where column amounts are high (Fig. 1b, d). Pennisi and LeCloarec's results²⁰ also show Cl:S ratios varying by a factor of ten in 24 h. We have not observed any such variability, possibly pointing to a greater inherent accuracy of the FTIR techniques.

Cospec measurements of the SO₂ emission rate from Etna were made simultaneously with our 14 June 1997 observations (63 kg s⁻¹, P. Allard, personal communication) and on 27 October 1997 (66 kg s⁻¹, T. Caltabiano, personal communication). Using these values, we obtain emission rates for HCl and HF of 8.9 kg s⁻¹

and 1.3 kg s⁻¹, respectively, for the June FTIR results, and 8.2 and 2.2 kg s⁻¹ for the October FTIR results. For HF, we regard the latter figure as more reliable because the HF:SO₂ ratio was better constrained at the proximal observation site (Fig. 1d). The SO₂ emission rate is typical of Etna's output for the last two decades, suggesting that the sustained HF emission rate is remarkably high: about 75 times the 0.029 kg s⁻¹ estimated for Kilauea³, and about 14% of the total anthropogenic HF output (Table 1). Etna's HF emissions are responsible for significant local environmental effects²².

Solar FTIR observations represent a major advance in techniques for monitoring volcanic gases. Not only does the Sun offer an intense 'free' source of energy to supplement those previously used, but the longer plume path lengths provide a greater signal for target gas absorption lines (provided they fall in atmospheric windows). Observations can be made up to a few tens of kilometres from the volcano and retrievals obtained in near real-time. In some cases, gases other than HCl, HF and SO₂ may also be detectable including OCS and CO (ref. 11). Observations are possible wherever the direct path between the Sun and instrument's field of view intersects the plume. We recently obtained solar occultation spectra of the plume from Soufriere Hills volcano, Montserrat, and found that useful data can be collected even at low solar elevations and with thin, high cloud partly occluding the Sun. Goff *et al.*²³ have recorded FTIR spectra of Popocatepetl volcano, Mexico, up to 10 km from the

Table 1 Comparative SO₂, HCl and HF emission rates; Cl:S and Cl:F mass ratios; and SO₂:HCl and HCl:HF molar ratios for Etna and other volcanoes worldwide

Volcano (source reference)	Cl:S mass ratio	SO ₂ :HCl molar ratio	Cl:F mass ratio	HCl:HF molar ratio	Emission rate (kg s ⁻¹)		
					SO ₂	HCl	HF
Etna, Italy							
14 June 1997	0.27	4.0	6.6	3.7	62.5*	8.9	1.3
26 October 1997 (this study)	0.24	4.6	3.83	2.1	66.2†	8.2	2.2
Etna‡							
July 1987 (ref. 18)	2.32	0.39	8.12	4.35	10.7	15.5	1.9
Kilauea, USA							
Summit§	0.006	66.5	n.d.	n.d.	1.62	0.05	0.03
Flank (ref. 3)	0.004	83.1	2.0	0.09	1.06	0.02	0.01
Erebus, Antarctica¶							
1986–91 (ref. 24)	1.4	0.79	3.8	1.1	0.49	0.35	0.17
Masaya 1980							
Nicaragua (ref. 25)	0.8	0.87	52	29.2	14.7	9.6	0.18
White Island,	2.4	2.5–2.6	51	24.6–25.7	3.7–4.0	0.8–0.9	0.017–0.02
New Zealand (ref. 26)							
Mt St Helens, USA#							
29 July 1980 (ref. 27)	0.33	5.83	2.1	2.0	14.8	1.44	0.37
Augustine, USA*							
1986/97 (ref. 5)	0.86	1.28	34	111	4.39	3.71	0.03
Global volcanism (refs 28, 29)					590	12.6–348	1.90–190
Anthropogenic (ref. 30)					~2,000	95.1	15.8

Etna, Kilauea, Erebus and Masaya are persistently degassing basaltic volcanoes. White Island is an andesitic volcano with persistent hydrothermal fumarolic activity. Mount St Helens and Augustine experienced explosive activity linked to extrusion of dacite lava domes and the emission rates given relate to brief eruptive episodes.

* Simultaneous aircraft measurements.

† Measured on 27 October 1997 by T. Caltabiano. The June Cospec results are more robust because of repetition of traverses and coincident timing with our FTIR measurements; the HF retrievals, however, are less accurate than those for October because concentrations were lower at the more distal observation site.

‡ All chloride calculated as HCl.

§ Average of four analyses. Emission rates for SO₂, HCl and HF calculated from S, Cl, and F budgets.

|| Average of six analyses. Emission rates for SO₂, HCl and HF calculated from S, Cl, and F budgets.

¶ Average of five analyses.

Non-explosive degassing.

* Typical non-explosive degassing. Average of three dome fumaroles.

summit, using sunlight scattered by high clouds and self-emission from target gases as infrared sources, extending further the array of configurations possible for remote gas measurement. There are some disadvantages, however. Thick cloud hinders observations, and not all volcanoes will emit SO₂, HCl and HF in sufficient quantities to exceed detection limits: these vary for individual gases and different combinations of field conditions. CO₂, an important volcanic gas, presents special difficulties because of its high ambient atmospheric concentrations¹³. Open-path FTIR retrievals of CO₂ appear to be most reliable over short path-lengths, using hot rocks or lamp sources¹¹.

Hitherto, our spectroscopic campaigns have been experimental and thus intermittent. Systematic observations maintained for longer periods at Etna and other volcanoes would permit detection of changes in HCl:SO₂ and HCl:HF ratios related to ascent, degassing and eruption of new magma^{7,20}, providing a new tool for volcano surveillance and eruption prediction. A sustained spectroscopic mission at Etna would also permit investigations of variations in chemistry between the various summit vents, and between these and flank vents, leading to improved models of its plumbing system. Lastly, solar occultation techniques open up the possibility of studying the atmospheric chemistry of tropospheric plumes because measurements can be made for variable plume age by moving up- or downwind, promising a clearer understanding of the scavenging of sulphur and halogens from plumes, and the formation of volcanic aerosol. □

Received 20 February; accepted 25 September 1998.

- Williams, S. N., Schaefer, S. J., Calvache, V. & Lopez, D. Global carbon dioxide emission to the atmosphere by volcanoes. *Geochim. Cosmochim. Acta* **56**, 1765–1770 (1992).
- McCormick, P. M., Thomason, L. W. & Trepte, C. R. Atmospheric effects of the Mt. Pinatubo eruption. *Nature* **373**, 399–404 (1995).
- Gerlach, T. M. & Graeber, E. J. Volatile budget of Kilauea volcano. *Nature* **313**, 273–277 (1985).
- Casadevall, T. J. et al. in *Volcanism in Hawaii* US Geol. Soc. Prof. Paper 1350, 771–780 (US Geol. Soc., Washington DC, 1987).

- Symonds, R. B., Rose, W. I., Gerlach, T. M., Briggs, P. H. & Harmon, R. S. Evaluation of gases, condensates and SO₂ emissions from Augustine volcano, Alaska: the degassing of a Cl-rich volcanic system. *Bull. Volcanol.* **52**, 355–374 (1990).
- Symonds, R. B., Rose, W. I., Bluth, G. J. S. & Gerlach, T. M. Volcanic-gas studies: methods, results and applications. *Rev. Mineral.* **30**, 1–66 (1994).
- Symonds, R. B., Mizutani, Y. & Briggs, P. H. J. Long term geochemical surveillance of fumaroles of Showa-Shinzan dome, Usu volcano, Japan. *Volcanol. Geotherm. Res.* **73**, 177–211 (1996).
- Stoiber, R. E., Malinconico, L. L. & Williams, S. N. in *Forecasting Volcanic Events* (eds Tazieff, H. & Sabroux, J.-C.) 425–444 (Elsevier, Amsterdam, 1980).
- Francis, P. W., Chaffin, C., Maciejewski, A. J. H. & Oppenheimer, C. M. M. Remote determination of SiF₄ in volcanic plumes: a new tool for volcano monitoring. *Geophys. Res. Lett.* **23**, 249–252 (1996).
- Francis, P. W., Maciejewski, A., Chaffin, C., Oppenheimer, C. & Caltabiano, T. SO₂ and HCl ratios in the plumes of Mt. Etna and Vulcano determined by Fourier transform spectroscopy. *Geophys. Res. Lett.* **22**, 1717–1720 (1995).
- Mori, T. & Notsu, K. Remote CO, COS, CO₂, SO₂ and HCl detection and temperature estimation of volcanic gas. *Geophys. Res. Lett.* **24**, 2047–2050 (1997).
- Hobbs, P. V., Radke, L. F., Lyons, J. H., Ferek, R. J. & Coffman, D. J. Airborne measurements of particle and gas emissions from the 1990 volcanic eruptions of Mount Redoubt. *J. Geophys. Res.* **96**, 18735–18752 (1991).
- Gerlach, T. M. et al. Application of the LI-COR CO₂ analyzer to volcanic plumes: a case study, volcano Popocatepetl, Mexico, June 7 & 10, 1995. *J. Geophys. Res.* **51**, 8005–8019 (1997).
- McGee, K. A. & Gerlach, T. M. Airborne volcanic plume measurements using an FTIR spectrometer, Kilauea volcano, Hawaii. *Geophys. Res. Lett.* **25**, 615–618 (1998).
- Bluth, G. J. S., Schnetzler, C. C., Kreuger, A. J. & Walter, L. S. The contribution of explosive volcanism to global atmospheric sulphur dioxide concentrations. *Nature* **366**, 327–329 (1993).
- Notholt, J., Toon, G. C., Lehmann, R., Sen, B. & Blavier, J. F. Comparison of arctic and antarctic trace gas column abundances from ground based FTIR spectrometry. *J. Geophys. Res.* **102**, 12863–12869 (1997).
- Allard, P., Carbonelle, J., Metrich, N. & Zettwoog, P. Eruptive and diffuse emissions of carbon dioxide from Etna volcano. *Nature* **351**, 387–391 (1991).
- Andres, R. J., Kyle, P. R. & Chuan, R. L. Sulphur dioxide, particle and elemental emissions from Mt. Etna, Italy, during July 1987. *Geol. Rundsch.* **82**, 687–695 (1993).
- Caltabiano, T., Romano, R. & Budetta, G. SO₂ flux measurements at Mt. Etna. *J. Geophys. Res.* **D6**, 12809–12819 (1994).
- Pennisi, M. & LeCloarec, M. Variations in Cl, F and S in Mt Etna's plume (Italy) between 1992 and 1995. *J. Geophys. Res.* **103**, 5061–5066 (1998).
- Le Guern, F. *Ecoulements Gazeux Réactifs à Hautes Températures, Mesures et Modélisation*. Thesis, Univ. Paris (1988).
- Notcutt, G. & Davies, F. B. N. Accumulation of volcanogenic fluoride by vegetation, Mount Etna, Sicily. *J. Volcanol. Geotherm. Res.* **39**, 329–333 (1989).
- Goff, F. et al. Geochemical surveillance of magmatic volatiles at Popocatepetl volcano, Mexico. *Geol. Soc. Am. Bull.* **110**, 695–710 (1998).
- Zreda-Gostyka, G., Kyle, P. R. & Finnegan, D. L. Chlorine, fluorine from Mt. Erebus, Antarctica, and estimated contributions to the Antarctic atmosphere. *Geophys. Res. Lett.* **20**, 1959–1962 (1993).
- Stoiber, R. E., Williams, S. & Huebert, B. J. Sulphur and halogen gases at Masaya caldera complex, Nicaragua: total flux and variations with time. *J. Geophys. Res.* **91**, 12,215–12,231 (1986).

26. Rose, W. I., Chaun, R. L., Giggenbach, W. F., Kyle, P. R. & Symonds, R. B. Rates of sulphur dioxide and particle emissions from White Island, New Zealand, and an estimate of the total flux of major gaseous species. *Bull. Volcanol.* **48**, 181–188 (1986).
27. McGee, K. A. The structure, dynamics and chemical composition of non-eruptive plumes from Mt. St. Helens 1980–88. *J. Volcanol. Geotherm. Res.* **51**, 269–282 (1992).
28. Stoiber, R., Williams, S. N. & Huebert, B. Annual contribution of sulphur dioxide to the atmosphere by volcanoes. *J. Volcanol. Geotherm. Res.* **33**, 1–8 (1987).
29. Symonds, R. B., Rose, W. I. & Reed, M. H. Contribution of Cl- and F-bearing gases to the atmosphere by volcanoes. *Nature* **334**, 415–418 (1988).
30. Cadle, R. D. A comparison of volcanic with other fluxes of atmospheric trace gas constituents. *Rev. Geophys.* **18**, 746–752 (1980).

Acknowledgements. We thank the EC Environment and Climate programme; the Open University, and UK Natural Environment research Council for funding; T. Caltabiano and the Istituto Internazionale per Vulcanologia, Catania for technical and logistic help in Italy; and the Parco dell'Etna for access to the volcano.

Correspondence and requests for materials should be addressed to P.F. (e-mail: p.w.francis@open.ac.uk).

Roots exert a strong influence on the temperature sensitivity of soil respiration

Richard D. Boone*, Knute J. Nadelhoffer†, Jana D. Canary* & Jason P. Kaye‡

* Institute of Arctic Biology, University of Alaska, Fairbanks, Alaska 99775, USA

† The Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA

‡ Department of Forest Science, Colorado State University, Ft Collins, Colorado 80523, USA

The temperature sensitivity of soil respiration will largely determine the effects of a warmer world on net carbon flux from soils to the atmosphere. CO₂ flux from soils to the atmosphere is estimated to be 50–70 petagrams of carbon per year and makes up 20–38% of annual inputs of carbon (in the form of CO₂) to the atmosphere from terrestrial and marine sources^{1,2}. Here we show that, for a mixed temperate forest, respiration by roots plus oxidation of rhizosphere carbon, which together produce a large portion of total effluxed soil CO₂, is more temperature-sensitive than the respiration of bulk soil. We determine that the Q₁₀ value (the coefficient for the exponential relationship between soil respiration and temperature, multiplied by ten) is 4.6 for autotrophic

root respiration plus rhizosphere decomposition, 2.5 for respiration by soil lacking roots and 3.5 for respiration by bulk soil. If plants in a higher-CO₂ atmosphere increase their allocation of photosynthate to roots^{3–6}, these findings suggest that soil respiration should be more sensitive to elevated temperatures, thus limiting carbon sequestration by soils.

The response of soil carbon fluxes to global warming is sensitive to slight changes in the relationship between soil temperature and soil respiration^{7,8}. Simulation models of regional and global carbon cycling generally use a single, fixed Q₁₀ coefficient for the exponential function between soil respiration and temperature^{9–11}. However, Q₁₀ varies among ecosystems and across temperature ranges, in part because the various components of soil respiration have different temperature sensitivities^{12–14}. These components include respiration by live roots (allowing their growth and maintenance) and associated mycorrhizae, and the oxidation of plant detritus (for example, roots, leaves and woody inputs), root exudates and humified organic matter by soil heterotrophs. We show here that variations in soil respiration through the growing season in a temperate hardwood forest are determined mainly by temperature responses of root respiration and rhizosphere heterotrophs (those in the area of soil immediately surrounding and influenced by plant roots). Microbial respiration outside the rhizosphere, although a significant fraction of total soil respiration, is less responsive to temperature than is root and rhizosphere respiration.

We used a long-term litter-manipulation experiment initiated in 1990–91 at the Harvard Forest, Petersham, Massachusetts, USA (42° 30' N, 72° 12' W), to examine the biological and physical controls on the dynamics of soil organic matter. The study site is an 85-year-old mixed-hardwood stand growing on a gentle, north-west-facing slope (4%) pastured in the 1800s. Soil is a well-drained stony, fine sandy loam, with a C_d horizon (hardpan) at 65 cm depth. Mean annual air temperature is 6°C; annual precipitation is 1,100 mm, distributed evenly through the year. Litter manipulations (3 m × 3 m plots, *n* = 3, except for control where *n* = 6) include 'control' (normal litter inputs), 'no litter' (above-ground litter excluded from plots annually), 'double litter' (above-ground litter doubled annually), 'no roots' (roots excluded from plots by fibre-glass-lined trenches), 'no inputs' (no above-ground litter and no roots), and 'OA-less' (organic (O) horizons and upper mineral soil (A) horizon (to 20 cm depth) removed and replaced with subsoil).

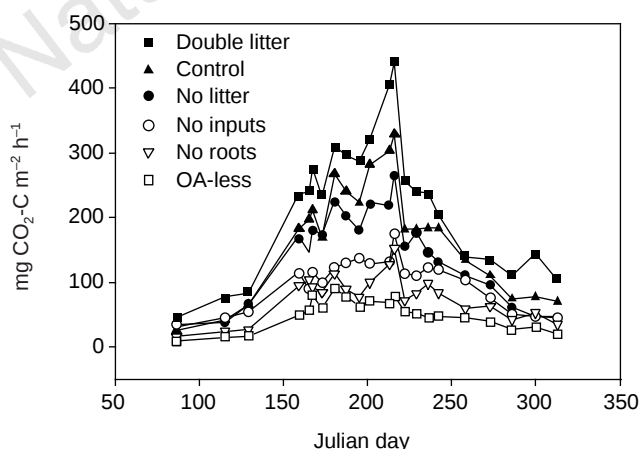


Figure 1 Soil CO₂ efflux per treatment for Harvard Forest litter-manipulation plots. Measurements were made over a one-year period from 16 June (Julian day 167) 1994 to 14 June (Julian day 165) 1995. Values are treatment means (*n* = 3 except for control, for which *n* = 6) of two measurements (early morning and late afternoon) within the same day. The means of coefficients of variation over the sampling period ranged from 10% to 15% by treatment. CO₂-C, carbon in the form of carbon dioxide.

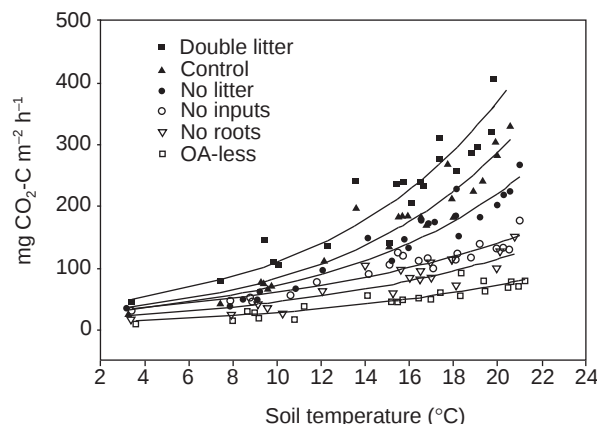


Figure 2 Relationship, for each treatment, between mean daily soil CO₂ flux and soil temperature at 5 cm depth. An exponential function of the form $y = \beta_0 e^{\beta_1 T}$, where y = CO₂ flux, β_0 and β_1 are fitted constants, and T = temperature, was applied to the data.

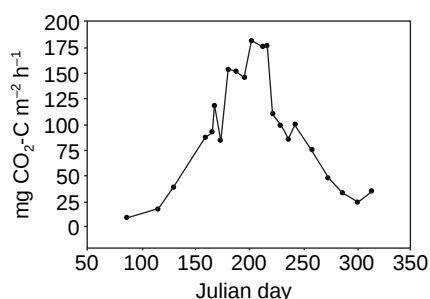


Figure 3 Soil CO₂ efflux by Julian day for 'roots', calculated as the difference in daily mean respiration between the control and no roots plots.

Measurements of soil respiration made in 1992 indicated that any increase in soil respiration due to decomposition of fine roots killed by the 1991 trenching of the 'no roots' and 'no inputs' plots was largely completed within one year of trenching¹⁵. We studied soil CO₂ efflux, temperature and moisture measured at the plots from June 1994 to June 1995.

Soil respiration (Fig. 1) varied by treatment, markedly tracked seasonal soil temperature (peaking in August for all treatments except OA-less) and was unrelated or weakly related to seasonal soil moisture. Soil temperature showed no treatment differences and ranged from 3.2 °C in late March to a maximum of 20.6 °C in early August, declining to 9.3 °C in late October. Soil moisture content (water volume in cm³ per soil volume in cm³) ranged from 0.21 to 0.54 for all treatments except OA-less and from 0.15 to 0.25 for OA-less; treatment means over the sampling period were 0.30–0.38 (all but OA-less) and 0.19 (OA-less). Soils generally were driest in July and wetter in spring and autumn. Soil moisture was always below saturation (~0.80 cm³ cm⁻³) and above the level (0.12 cm³ cm⁻³) identified previously for similar Harvard Forest soils as the breakpoint for a water limitation on respiration¹⁶. Multiple regression analysis using soil temperature and soil moisture as independent variables showed that soil water content (relatively constant during the measurement period) was unrelated to soil respiration for the litter treatments and was weakly related ($R^2 = 0.86$, temperature alone; $R^2 = 0.90$, temperature plus moisture) for the control plots.

We applied an exponential curve of the form $y = \beta_0 e^{\beta_1 T}$, where y is the carbon flux, β_0 and β_1 are fitted constants and T is the temperature, to compare the relationship between soil respiration and soil temperature (Fig. 2) and calculated Q_{10} values, where $Q_{10} = e^{10 \times \beta_1}$. We calculated 'roots' respiration (Fig. 3) as the difference in daily mean respiration rates between the control and no roots plots, and derived a 'roots' Q_{10} value (Fig. 4). Within each treatment there was a strong correlation between temperature and respiration, but the sensitivity of the relationship (indicated by Q_{10} values and regression slopes) varied among treatments (Fig. 2 and Table 1). Q_{10} values and slopes changed nonsignificantly with either addition or exclusion of leaf litter, whereas treatments without roots had lower Q_{10} ($P = 0.053$, no roots; $P < 0.05$, no inputs) and slopes ($P < 0.05$) than those for the control. The 'roots' Q_{10} value and

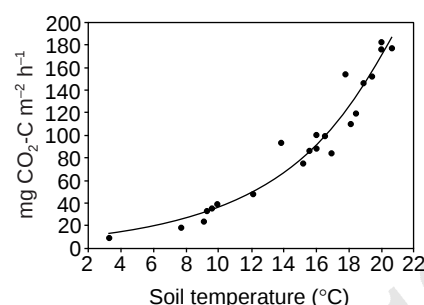


Figure 4 Relationship between mean daily CO₂ flux (calculated) from 'roots' and soil temperature at 5 cm depth. Soil temperatures are the means of control and no roots plots. An exponential function (Fig. 2) was applied to the data.

slope were significantly greater ($P < 0.05$) than those for the control and the treatments without roots. The Q_{10} value for the control plots (3.5) was close to the values for soil respiration determined elsewhere at the Harvard Forest ($Q_{10} = 3.9$; ref. 16) and for hardwood forests globally ($Q_{10} = 3.1$; ref. 10). The Q_{10} values for plots without roots are within the limits reported for soils (without roots) incubated in the laboratory over a similar temperature range^{17,18}.

Our Q_{10} value for 'roots' (4.6) is much higher than that reported previously for root respiration^{19–23}. However, there is an important difference between our approach and others. Previous measurements for root Q_{10} values have been for autotrophic root respiration alone, determined by assays of roots (often excised) that were free of *in situ* rhizosphere soil. Our Q_{10} values reflect not only root respiration but also respiration associated with the roots, including respiration by mycorrhizae and the decomposition of labile root-derived organic material (detritus and exudates) by microbiota in the rhizosphere. Increased production of root exudates (perhaps due to higher membrane permeability) at higher temperatures may be another contributing factor. If root respiration *per se* at our site shows typical temperature sensitivity (for example, Q_{10} values in the range 2–3), the temperature sensitivity and Q_{10} values for the mycorrhizae and rhizosphere heterotrophs together must be much higher than 2–3.

The seasonal pattern in soil respiration probably resulted in part from changes in root biomass and production. However, such changes would confound results only if they are unrelated to soil temperature. The synchronicity in soil respiration in the control and no roots plots indicates that root phenology is probably not independent of temperature.

The higher Q_{10} value for roots plus associated rhizosphere has implications for analysis of how soil CO₂ efflux may be affected in a warmer world. The results indicate that the temperature sensitivity of soil respiration may depend on the relative contribution that roots and associated rhizosphere microbiota (including mycorrhizae) make to total soil CO₂ efflux. Systems most sensitive to temperature rise should be those in which roots and the associated rhizosphere contribute the largest portion of total CO₂ flux. We would also expect that soil Q_{10} values should rise if higher atmospheric CO₂ were to lead to higher below-ground carbon allocation^{3–6}. The factors that determine the relative contribution of roots and rhizosphere to CO₂ flux (this contribution can range from 20% to 90% of the total flux^{15,24–28}) have not been well evaluated. The contribution of rhizosphere versus non-rhizosphere microorganisms to soil respiration is also poorly understood. More information across ecosystem types on the temperature sensitivities of soil CO₂ sources and their relative contributions to total soil respiration, and on how global change may influence below-ground photosynthate allocation, is essential to allow us to predict how increased temperature affects soil CO₂ efflux. □

Methods

Soil CO₂ efflux, temperature and moisture were measured at the plots from

Table 1 R^2 and Q_{10} values for the relationship between soil respiration and temperature

Treatment	R^2	Q_{10}
Control	0.91	3.5 (0.4)
Double litter	0.90	3.4 (0.4)
No litter	0.91	3.1 (0.3)
No roots	0.73	2.5 (0.4)
No inputs	0.89	2.3 (0.2)
OA-less	0.82	2.6 (0.3)
'Roots'	0.95	4.6 (0.5)

For R^2 values; $P < 0.01$; Q_{10} values are means ($\pm$ s.e.m.). Q_{10} values were obtained from the exponential curve of the form $y = \beta_0 e^{\beta_1 T}$, where $Q_{10} = e^{10 \times \beta_1}$. Standard error for Q_{10} is calculated as $Q_{10} \times 10 \times \text{s.e.}(\beta_1)$.

June 1994 to June 1995. Soil CO₂ efflux was determined with an infrared gas analyser (IRGA, LICOR Model 6262) and a flow-through chamber. CO₂ concentrations were recorded every 6 s over a 4-min period in a tube-shaped chamber (0.25 m diameter × 0.10 m height) placed on a permanent collar inserted 1 cm into the soil; air was circulated by rotary pump through the system at 0.85 l min⁻¹. The chamber was vented with a capillary tube to allow equilibration of air pressure. Measurements were taken at sunrise during minimum flux and in the afternoon during maximum flux, as determined by previous diel sampling. Soil respiration rates reported here are the means of the two measurements for each treatment. Soil moisture at the time of the CO₂ measurements was determined using time domain reflectometry (TDR) probes (0–15 cm depth) positioned vertically through the forest floor and mineral soil. Soil dielectric values for all but the OA-less plots (no forest floor) were converted to volumetric water content on the basis of a calibration derived for nearby similar soils²⁹; a calibration for mineral soil³⁰ was used for the OA-less plots. Soil temperature was measured hourly using Campbell 107-B temperature thermistors (5 cm depth); those reported are treatment averages of 24 measurements taken hourly on the day of the CO₂ flux measurements. CO₂ flux measurements were made weekly from June to August, biweekly in September and October, and monthly in all other months except December, January and February, when no measurements were taken. Q₁₀ values and regression slopes of respiration (log₁₀-transformed) versus temperature were compared by Student's *t*-test.

Received 11 August; accepted 14 October 1998.

1. Raich, J. W. & Schlesinger, W. H. The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. *Tellus* **44B**, 81–99 (1992).
2. Raich, J. W. & Potter, C. S. Global patterns of carbon dioxide emissions from soils. *Global Biogeochem. Cycles* **9**, 23–26 (1995).
3. Fitter, A. H. *et al.* Root production and mortality under elevated atmospheric carbon dioxide. *Plant Soil* **187**, 299–306 (1996).
4. Canadell, J. G., Pitelka, L. F. & Ingram, J. S. I. The effects of elevated [CO₂] on plant-soil carbon background: a summary and synthesis. *Plant Soil* **187**, 391–400 (1997).
5. Cardon, Z. Influence of rhizodeposition under elevated CO₂ on plant nutrition and soil organic matter. *Plant Soil* **187**, 277–288 (1997).
6. Hungate, B. A. *et al.* The fate of carbon in grasslands under carbon dioxide enrichment. *Nature* **388**, 576–579 (1997).
7. Holland, E. A., Townsend, A. R. & Vitousek, P. M. Variability in temperature regulation of CO₂ fluxes and N mineralization from five Hawaiian soils: implications for a changing climate. *Global Change Biol.* **1**, 115–123 (1995).
8. Townsend, A. R., Vitousek, P. M. & Holland, E. A. Tropical soils dominate the short term carbon cycle feedbacks to atmospheric carbon dioxide. *Clim. Change* **22**, 293–303 (1992).
9. Fung, I. Y., Tucker, C. J. & Prentice, K. C. Application of advanced very high resolution vegetation index to study atmosphere-biosphere exchange of CO₂. *J. Geophys. Res.* **92**, 2999–3015 (1987).
10. Kicklighter, D. W. *et al.* Aspects of spatial and temporal aggregation in estimating regional carbon dioxide fluxes from temperate forest soils. *J. Geophys. Res.* **99**, 1303–1315 (1994).
11. Xiao, X. *et al.* Transient climate change and net ecosystem production of the terrestrial biosphere. *Global Biogeochem. Cycles* **12**, 345–360 (1998).
12. Schleser, G. H. The response of CO₂ evolution from soils to global temperature changes. *Z. Naturforsch.* **37**, 287–291 (1982).
13. Kirschbaum, M. U. F. The temperature dependence of soil organic matter decomposition, and the effect of global warming on soil organic C storage. *Soil Biol. Biochem.* **27**, 753–760 (1995).
14. Trumbore, S. E., Chadwick, O. A. & Amundson, R. Rapid exchange between soil carbon and atmospheric carbon dioxide driven by temperature change. *Science* **272**, 393–396 (1996).
15. Bowden, R. D., Nadelhoffer, K. J., Boone, R. D., Melillo, J. M. & Garrison, J. B. Contributions of aboveground litter, belowground litter, and root respiration to total soil respiration in a temperate mixed hardwood forest. *Can. J. For. Res.* **23**, 1402–1407 (1993).
16. Davidson, E. A., Belk, E. & Boone, R. D. Soil water content and temperature as independent or confounded factors controlling soil respiration in a temperate mixed hardwood forest. *Global Change Biol.* **4**, 217–228 (1998).
17. Ross, D. J. & Tate, K. R. Microbial C and N, and respiratory activity, in litter and soil of a southern beech (*Nothofagus*) forest: distribution and properties. *Soil Biol. Biochem.* **25**, 477–483 (1993).
18. Winkler, J. P., Cherry, R. S. & Schlesinger, W. H. The Q₁₀ relationship of microbial respiration in a temperate forest soil. *Soil Biol. Biochem.* **28**, 1067–1072 (1996).
19. Lawrence, W. T. & Oechel, W. C. Effects of soil temperature on the carbon exchange of taiga seedlings. I. Root respiration. *Can. J. For. Res.* **13**, 840–849 (1983).
20. Burton, A. J., Pregitzer, K. S., Zogg, G. P. & Zak, D. R. Latitudinal variation in sugar maple fine root respiration. *Can. J. For. Res.* **26**, 1761–1768 (1996).
21. Ryan, M. G., Hubbard, R. M., Pongracic, S., Raison, R. J. & McMurtrie, R. E. Foliage, fine-root, woody tissue and stand respiration in *Pinus radiata* in relation to nitrogen status. *Tree Physiol.* **9**, 255–266 (1996).
22. Zogg, G. P., Zak, D. R., Burton, A. J. & Pregitzer, K. S. Fine root respiration in northern hardwood forests in relation to temperature and nitrogen availability. *Tree Physiol.* **16**, 719–725 (1996).
23. Burton, A. J., Pregitzer, K. S., Zogg, G. P. & Zak, D. R. Drought reduces root respiration in sugar maple forests. *Ecol. Appl.* **8**, 771–773 (1998).
24. Anderson, J. M. Carbon dioxide evolution from two temperate, deciduous woodland soils. *J. Appl. Ecol.* **10**, 361–378 (1973).
25. Dörr, H. & Münnich, K. O. Annual variation in soil respiration in selected areas of the temperate zone. *Tellus* **39B**, 114–121 (1987).
26. Johnson, D. *et al.* Soil pCO₂, soil respiration, and root activity in CO₂-fumigated and nitrogen-fertilized ponderosa pine. *Plant Soil* **165**, 129–138 (1994).
27. Thierion, V. & Laudelout, H. Contribution of root respiration to total CO₂ efflux from the soil of a deciduous forest. *Can. J. For. Res.* **26**, 1142–1148 (1996).
28. Hanson, P. J., Edwards, N. T., Garten, C. T. Jr & Andrews, J. A. Separating root and soil microbial

- contributions to soil respiration: a review of methods and observations. *Biogeochemistry* (submitted).
29. Newkirk, K. M., Bowles, F. P. & Catricala, C. E. Time domain reflectometry in forest soils: *in-situ* use at the Harvard Forest Soil Warming Experiment. *Agron. Abstr.* 235 (1994).
 30. Topp, G. C., Davis, J. L. & Annan, A. P. Electromagnetic determination of soil water content: measurements in coaxial transmission lines. *Wat. Resour. Res.* **16**, 574–582 (1980).

Acknowledgements. We thank E. Davidson and E. Sundquist for advice on construction of the field system for measuring soil respiration; E. Davidson for help with analysis of soil respiration data; K. Newkirk and F. Bowles for help with the design, fabrication and use of the TDR system; and A. Doyle for critical review of the manuscript. This work was sponsored by grants from the US NSF Long-Term Ecological Research Program, the Department of Energy NIGEC Program and the US Department of Agriculture Competitive Grants Program.

Correspondence and requests for materials should be addressed to R.D.B. (e-mail: ffrdb@uaf.edu).

Selective sweep of a newly evolved sperm-specific gene in *Drosophila*

Dmitry I. Nurminsky*, Maria V. Nurminskaya†, Daniel De Aguiar*, Daniel L. Hartl*

* Harvard University, Department of Organismic & Evolutionary Biology, Cambridge, Massachusetts 02138, USA

† Tufts University School of Medicine, Department of Anatomy and Cell Biology, Boston, Massachusetts 02111, USA

The pattern of genetic variation across the genome of *Drosophila melanogaster* is consistent with the occurrence of frequent 'selective sweeps', in which new favourable mutations become incorporated into the species so quickly that linked alleles can 'hitchhike' and also become fixed¹. Because of the hitchhiking of linked genes, it is generally difficult to identify the target of any putative selective sweep. Here, however, we identify a new gene in *D. melanogaster* that codes for a sperm-specific axonemal dynein subunit. The gene has a new testes-specific promoter derived from a protein-coding region in a gene encoding the cell-adhesion protein annexin X (*AnnX*), and it contains a new protein-coding exon derived from an intron in a gene encoding a cytoplasmic dynein intermediate chain (*Cdic*). The new transcription unit, designated *Sdic* (for sperm-specific dynein intermediate chain), has been duplicated about tenfold in a tandem array. Consistent with the selective sweep of this gene, the level of genetic polymorphism near *Sdic* is unusually low. The discovery of this gene supports other results that point to the rapid molecular evolution of male reproductive functions^{2–4}.

Our initial observation was that the genetic organization of region 19DE on the X chromosome of *D. melanogaster* differs from that of other species in the *melanogaster* subgroup^{5,6}. The *D. melanogaster* genome contains an additional ~70 kilobases (kb) of DNA, consisting of ~10 tandem repeats of a unit ~7 kb in length. The tandem repeat is flanked at its 5' end by *Cdic*, which encodes the cytoplasmic dynein intermediate chain, and at its 3' end by *AnnX*, which encodes annexin X. The repeating unit is formed from a fusion of the central region of *AnnX* with the 3' region of *Cdic*. One possible scenario (Fig. 1) is that the *Cdic*–*AnnX* region became tandemly duplicated; then, a deletion fused *AnnX* exon a4 with *Cdic* intron 3, another deletion eliminated the 5' end of *AnnX* extending into exon a2, and a third deletion with breakpoints in *Cdic* introns 5 and 7 eliminated exon v2. This new structure was then tandemly duplicated. We designated the repeating unit *Sdic* because subsequent studies showed it to encode a sperm-specific axonemal dynein intermediate chain.

Nucleotide sequence analysis of complementary DNA clones gave unambiguous evidence for transcription of the *Sdic* unit⁶. Only the *Cdic*-derived sequences in *Sdic* are transcribed, and the *Sdic* messenger RNA contains four of the five *Cdic* exons present in the repeat. Exon v1 is not included; this exon is alternatively spliced in

Cdic and present in some but not all *Cdic* mRNAs. The exons in *Sdic* code for a protein of approximate relative molecular mass 60,000 ($M_r \sim 60K$) derived from the carboxy-terminal portion of the dynein intermediate chain molecule. This region of the intermediate chain is apparently responsible for the interaction of the chain with other components of the dynein complex, and its sequence is strongly conserved between cytoplasmic and axonemal dynein intermediate chains. It seemed unlikely that the *Sdic* product could function as a subunit of cytoplasmic dynein, as it lacks the first three exons of *Cdic*. These exons code for an amino-terminal coiled-coil domain as well as a serine-rich domain, both of which are essential for the function of cytoplasmic dynein intermediate chains⁷. *Sdic* also contains an unusual 5' exon, which encodes the N-terminal 16 amino acids; this exon is derived from sequence present in intron 3 of the *Cdic* gene. Comparison of this acquired N-terminal domain with the corresponding regions of the intermediate chains of axonemal dyneins revealed 33% identity and 57% similarity across the initial 21 residues (Fig. 2). These regions also have similar numbers of residues between the N-terminal amino acid and the conserved PPE/TQT motif⁶ used as a landmark in the alignment (Fig. 2).

The protein comparisons indicate that, if *Sdic* is functional, its product might be an axonemal dynein intermediate chain. In *Drosophila*, the axoneme constitutes a major part of the sperm tail. We found that transcription of *Sdic* is not only abundant in testes but is also testes-specific (Fig. 3). We created an *Sdic::GFP*

reporter cassette, coding for the *Sdic* polypeptide fused at its C-terminal end to GFP (green fluorescent protein), under the control of the *Sdic* promoter. Transgenic flies carrying this cassette expressed *Sdic::GFP* specifically in spermatocytes (Fig. 4). The *Sdic::GFP* fusion protein accumulated in mature spermatocytes and ultimately became incorporated into the sperm tails. The localization of *Sdic::GFP* is consistent with *Sdic* encoding a new axonemal dynein intermediate chain.

Acquisition of the *Sdic* function also required the evolution of testes-specific gene regulation. Neither the *AnnX* nor the *Cdic* promoter regions are present in the *Sdic* repeat. The 5' ends of the *Sdic* mRNAs map immediately downstream of the junction between *AnnX* exon a4 and *Cdic* intron 3. This junction created a new promoter with extensive similarity to the *Cdic* promoter, but with a completely independent origin (Fig. 5). A proximal element of 18 base pairs (bp) shares 80% identity with the proximal region of the *Cdic* promoter, although the former derives from an intron (intron 3 of *Cdic*). A distal element of 32 bp shares 74% identity with the distal region of the *Cdic* promoter, although the former derives from exon a4 of *AnnX*. A database search also revealed a region of *Sdic* with high similarity to a testis-specific promoter element (TSE) previously characterized in the testes-specific $\beta 2$ -tubulin gene⁸. This region of 27 bp shares 72% identity with the TSE (Fig. 5), but it derives from the junction between exon a4 of *AnnX* and intron 3 of *Cdic*. The new promoter present in *Sdic* can account for the testes-specific transcription of the newly evolved gene.

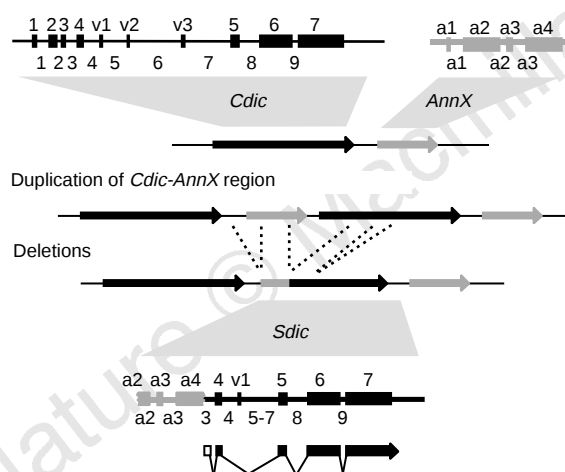


Figure 1 A model for the creation of the *Sdic* gene. Top, exon-intron structure of the *Cdic* and *AnnX* genes. Exons are numbered above the line and introns below the line; v denotes a variable (alternatively spliced) exon. The model below postulates a duplication of the *Cdic* (black arrow)-*AnnX* (grey arrow) region, followed by three deletions (dashed lines) to create *Sdic*. The *Sdic* gene has also been tandemly duplicated ~10-fold. At the bottom is the transcription and splicing pattern of the *Sdic* units.

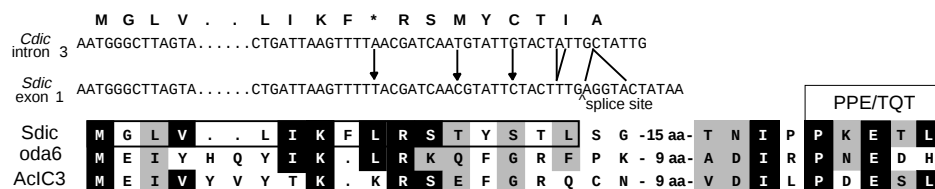
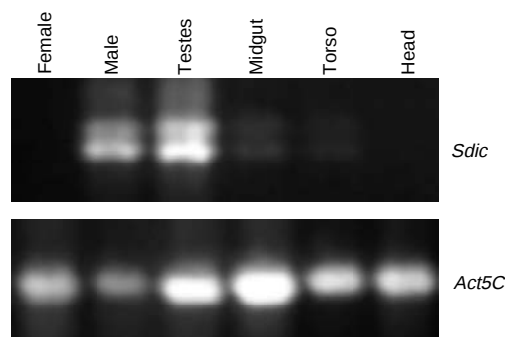


Figure 2 Creation of the first exon of *Sdic* from the sequence of intron 3 of *Cdic*, and a comparison of the N-terminal domain of *Sdic* coded by this first exon with the N-terminal domains of axonemal dynein intermediate chains from *Chlamydomonas reinhardtii* (oda6, GenBank accession number X55382) and *Anthocidaris crassipina* (AclC3, GenBank accession number D28863). Theoretical products of translation of the original *Cdic* intron sequence are shown at

Figure 3 Testis specificity of *Sdic* transcription, demonstrated by reverse transcription and the polymerase chain reaction. The first cDNA strand was synthesized from mRNA isolated from adult females, males or male body parts, as shown at the top. *Sdic* cDNA was successfully amplified from males (specifically from the testes) but not females. Control amplification of the cDNA encoding a constitutively expressed actin gene (*Act5C*, GenBank accession number K00668) shows that there is enough template in all six samples.



the top; the asterisk denotes a stop codon. Nucleotide-sequence differences between the *Cdic* intron and the *Sdic* exon are shown by arrows. The new donor splice site created by an 11-bp insertion is indicated. In the amino-acid alignment, identical residues are shown in black boxes and similar residues in grey boxes. Gaps in the alignment are indicated by dots. The PPE/TQT motif is conserved in both axonemal and cytoplasmic dynein intermediate chains⁸.

The testes-specific promoter seems to have been created largely through serendipity by the juxtaposition of sequences that strongly resemble promoter elements, even though they originated from non-regulatory sequences. The 18-bp proximal promoter element differs by only 1 bp from the homologous *Cdic* intron 3 sequence, and the 32-bp distal promoter element is identical in sequence to the homologous *AnnX* exon (Fig. 5). Most interesting is the TSE formed by the *AnnX*–*Cdic* junction, which differs from homologous *AnnX* sequence by 5 of 27 base pairs (Fig. 5).

In contrast to the *Sdic* promoter, the coding sequence required substantial modification to create the new 5' exon coding for the 16 N-terminal residues (Fig. 2). This coding region derives entirely from *Cdic* intron 3. The homologous sequence in *Cdic* intron 3 is untranslatable. In the same reading frame as *Sdic*, it includes a nonsense codon within the first 48 nucleotides. The most parsimonious alignment of the sequences indicates that formation of the new coding sequence must have required at least three nucleotide substitutions, one single-nucleotide deletion, and an 11-bp insertion creating the new donor splice site (Fig. 2). As the *Sdic* promoter appears to have acquired testes specificity of expression almost instantaneously, it seems likely that the selection pressure for the new dynein function occurred in the testes.

What suggests that *Sdic* is the result of a selective sweep? First, there has been extensive refashioning of part of *Cdic* intron 3 to form a new coding region with the sequence characteristics and proper length to serve as the N-terminal end of an axonemal dynein intermediate chain. Second, it is unlikely that the ~10-fold repeat of *Sdic*, comprising ~70 kb of DNA, would have originated and been maintained in the absence of positive selection, in view of the high rate of spontaneous deletion of unconstrained sequences in the *Drosophila* genome⁹. Third, all *D. melanogaster* strains studied contain *Sdic*, whereas other members of the species subgroup, even closely related species, do not (data not

shown). Finally, the present level of genetic variation in 1,200 bp of *Sdic* and 985 bp of *Cdic* was determined in each of nine *D. melanogaster* strains of geographically diverse origin; these data yield estimates of nucleotide polymorphism (θ) of $1.23 \times 10^{-3} \pm 0.83 \times 10^{-3}$ (*Sdic*) and $0.78 \times 10^{-3} \pm 0.66 \times 10^{-3}$ (*Cdic*), and estimates of nucleotide diversity (π) of $0.89 \times 10^{-3} \pm 0.73 \times 10^{-3}$ (*Sdic*) and $0.45 \times 10^{-3} \pm 0.50 \times 10^{-3}$ (*Cdic*). These are among the smallest values of polymorphism and diversity found in nuclear genes of diverse geographic isolates of *Drosophila*¹⁰.

The data are consistent with a relatively recent selective sweep in the *Sdic* region, either of *Sdic* itself or of a closely linked gene. The time of origin of *Sdic* is unclear. It could trace back to the time of speciation, about 3 million years ago¹¹. *Sdic* might even have been involved in speciation, perhaps through reduced male fertility or sterility in nascent species hybrids, but we have no evidence bearing on this possibility. Incidentally, the homologous coding regions of *Sdic* and *Cdic* include six non-synonymous and two synonymous nucleotide substitutions, whereas the present polymorphisms of *Sdic* include two non-synonymous and two synonymous changes. Although these numbers are too small to prove anything, they suggest a pattern that would be expected if positive selection were governing the fixed differences and if stabilizing selection were governing current polymorphisms. This is the pattern observed for *jingwei*, an apparently newly evolved gene with an unidentified function in *D. teissieri* and *D. yakuba*¹².

The evolution of *Sdic* contrasts with several widely held views about the evolution of new genetic functions. First, although a testes-specific promoter was essential for *Sdic*, this unusual regulatory region did not really 'evolve'. Instead it was aboriginal, created *de novo* by the fortuitous juxtaposition of suitable sequences. The more extensive evolutionary changes took place in *Cdic* intron 3, enabling an originally untranslatable sequence to become a new coding region whose product functions in the assembly of axonemal

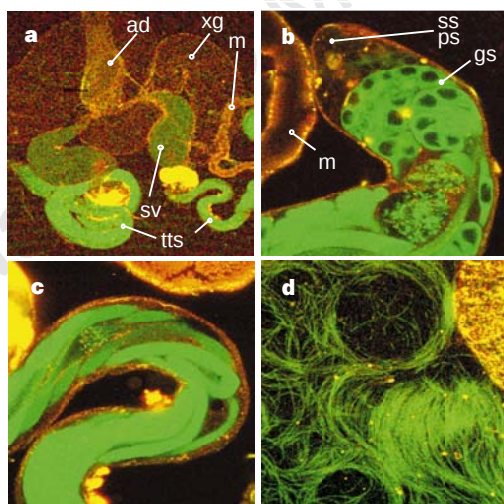


Figure 4 Testes distribution of an *Sdic*::GFP fusion protein, whose translation is driven by the *Sdic* promoter. **a**, The reproductive system of a transgenic male carrying the fusion construct. The *Sdic*::GFP protein (green) is expressed in testes (tts) and is present in the sperm vesicles (sv), but it is not present in accessory glands (xg) or the anterior ejaculatory duct (ad). The midgut (m) also shows no expression. **b**, The *Sdic*::GFP fusion protein is not present in the stem cells (ss) and proliferating spermatocytes (ps), but does appear in growing spermatocytes (gs). **c**, *Sdic*::GFP is abundant in the bundles of maturing spermatocytes. **d**, *Sdic*::GFP strongly marks the tails of mature sperm released into the surrounding medium.

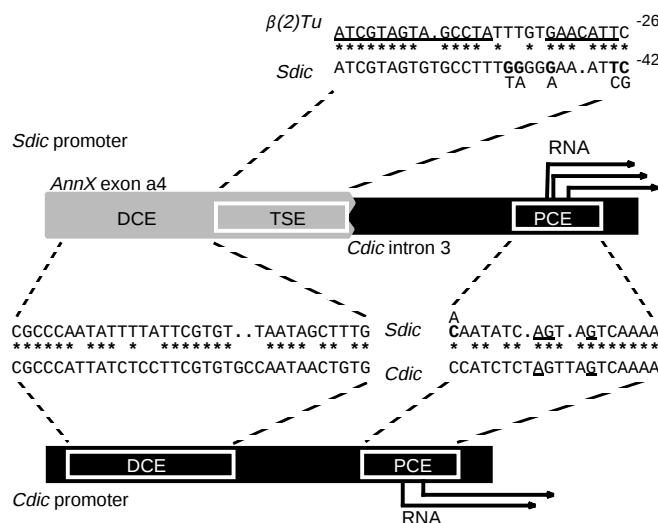


Figure 5 Detailed structure of the *Sdic* promoter formed by the junction between *AnnX* exon a4 and *Cdic* intron 3. Nucleotide alignments are presented for regions of similarity between the *Sdic* and *Cdic* promoters, designated DCE (distal conserved element) and PCE (proximal conserved element). Identical positions are indicated by asterisks and gaps in the alignment by dots. In the local alignment for PCE, nucleotides corresponding to the RNA start sites are underlined; nucleotides differing from the homologous *Cdic* intron sequence are in bold type, with the *Cdic* nucleotides shown on top. The DCE sequence in the *Sdic* promoter is identical to the homologous sequence in *AnnX* exon a4. The testes-specific element (TSE) is also shown and aligned with the TSE from the β 2-tubulin gene ($\beta(2)Tu$). Nucleotides essential for TSE function in the β 2-tubulin gene are underlined⁸. Differences from the homologous sequence in *AnnX* exon a4 are shown in bold type, with the *AnnX* nucleotides shown at the bottom.

dynein. Second, in the evolution of the *Sdic* gene, there has been no conserved distinction between a 'regulatory' sequence and a 'coding' sequence. The *Sdic* promoter was created in large part from pre-existing coding sequences in *AnnX* exon a4, and the new *Sdic* N-terminal end evolved from *Cdic* intron 3 sequences with no pre-existing coding function. Third, there is a high degree of identity between the proximal and distal promoter elements of *Sdic* and those of *Cdic*, between the TSE promoter element of *Sdic* and that of the testes-specific $\beta 2$ -tubulin gene, and between the amino-acid sequences of the N-terminal ends of *Sdic* and other axonemal dynein intermediate chains. Yet these similar sequence elements have completely different origins. Without knowing the history of *Sdic*, it is tempting to attribute the close similarities to common ancestry and 'exon shuffling'¹³. We do not yet know how *Sdic* contributes to the function of the sperm axoneme, or even whether it is essential for male fertility. The rapid evolution of the *Sdic* gene does, however, support evidence from interspecific male sterility³, as well as from direct molecular comparisons^{2,4}, that genes involved in male reproductive functions may evolve faster than other types of genes. □

Methods

Transcription analysis. We isolated total RNA from adult flies or body parts obtained by hand dissection, and purified poly(A)⁺ RNA¹⁴. Reverse transcription was performed using oligo(T)-adaptor primer 5'-ACTGCTCGTACCAC AATGGCTGCT-3' and Superscript reverse transcriptase (Gibco). Full-length *Sdic* coding sequences were amplified by the polymerase chain reaction in thermostable polymerase mix (Elongase, Gibco) using primers *Sdic*-F (5'-GG TACCTATTACACAGATTGGACAATGGGC-3') and *Cdic*-LL (5'-GGCGGCC GCGTTCATCTTGATCTCGCTAAG-3'). Primer extension experiments were performed on poly(A)⁺ RNA from adult males, using primers *Sdic*-P-L (5'-ATTGTCGAATCTGTGTAATA-3') for *Sdic* and *Cdic*-P-L (5'-CACCTCTTGACTGTCTTTACT-3') for *Cdic*. Cloning of the *Sdic* cDNA is described in ref. 6.

Plasmid constructs. A reporter construct *SdicP/Sdic::GFP/SV40* was made using the vector for P-element-mediated transformation pYES¹⁵, and the modified GFP-coding sequence was from vector pGreenLantern (Gibco). We amplified a 390-bp fragment containing the *Sdic* promoter using primers *Sdic*-P-U (5'-TCGGGCAAAGTAGTAATTGTA-3') and *Sdic*-P-L, and cloned in pCRII. The *Sdic* promoter fragment was excised by *Xho*I and *Kpn*I, and was combined in pSP72 (Promega) with the *Kpn*I-*Nof*I fragment containing the *Sdic* ORF, and also with the *Nof*I-*Cl*aI fragment of pGreenLantern carrying the GFP ORF and the SV40 transcriptional terminator. The resulting cassette, comprising *SdicP* promoter/*Sdic::GFP* fusion ORF/SV40 terminator, was excised by digestion with *Xho*I and *Bgl*II. The *Bgl*II end was rendered blunt with Klenow fragment, and the cassette inserted into pYES previously digested with *Xho*I and *Sma*I.

Germline transformation. DNA of *SdicP/Sdic::GFP/SV40* in pYES was purified in a column matrix (Quiagen), mixed with the helper plasmid p π 25.7wc (ref. 16), and injected into embryos of a *D. melanogaster* *y w* strain as described¹⁷. Transgenic flies were selected for the *y*⁺ phenotype.

Fluorescent microscopy. Transgenic males were dissected in 0.15 M NaCl and their reproductive system was observed in an Axiovert 100TV laser scanning microscope (Zeiss). GFP fluorescence was detected in the fluorescein isothiocyanate channel and red autofluorescence of the tissues was detected in the rhodamine channel. Images were processed with Adobe Photoshop.

Received 1 September; accepted 26 October 1998.

- Begun, D. J. & Aquadro, C. F. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356**, 519–520 (1992).
- Coulthart, M. B. & Singh, R. S. High level of divergence of male-reproductive-tract proteins, between *Drosophila melanogaster* and its sibling species, *D. simulans*. *Mol. Biol. Evol.* **5**, 182–191 (1988).
- Wu, C.-I. & Davis, A. W. Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic basis. *Am. Nat.* **142**, 187–212 (1993).
- Civetta, A. & Singh, R. S. Sex-related genes, directional sexual selection, and speciation. *Mol. Biol. Evol.* **15**, 901–909 (1998).
- Benevolenskaya, E. V., Nurminsky, D. I. & Gvozdev, V. A. Structure of the *Drosophila melanogaster* annexin X gene. *DNA Cell Biol.* **14**, 349–357 (1994).
- Nurminsky, D. I. *et al.* Cytoplasmic dynein intermediate chain isoforms with different targeting properties created by tissue-specific alternative splicing. *Mol. Cell. Biol.* **18**, 6816–6825 (1998).
- Steffen, W. *et al.* The involvement of the intermediate chain of cytoplasmic dynein in binding the motor complex to membranous organelles of *Xenopus* oocytes. *Mol. Biol. Cell* **8**, 2077–2088 (1997).
- Michiels, F., Gasch, A., Kaltschmidt, B. & Renkawitz-Pohl, R. A 14 bp promoter element directs the

testis specificity of the *Drosophila beta 2 tubulin* gene. *EMBO J.* **8**, 1559–1565 (1989).

- Petrov, D. A., Lozovskaya, E. R. & Hartl, D. L. High intrinsic rate of DNA loss in *Drosophila*. *Nature* **384**, 346–349 (1996).
- Moriyama, E. N. & Powell, J. R. Intraspecific nuclear DNA variation in *Drosophila*. *Mol. Biol. Evol.* **13**, 261–277 (1996).
- Caccone, A., Amato, G. D. & Powell, J. R. Rates and patterns of scnDNA and mtDNA divergence within the *Drosophila melanogaster* subgroup. *Genetics* **118**, 671–683 (1988).
- Long, M. & Langley, C. H. Natural selection and the origin of *jingwei*, a chimeric processed functional gene in *Drosophila*. *Science* **260**, 91–95 (1993).
- Long, M., Rosenberg, C. & Gilbert, W. Intron phase correlations and the evolution of the intron/exon structure of genes. *Proc. Natl Acad. Sci. USA* **92**, 12495–12499 (1995).
- Nurminsky, D. I., Moriyama, E. N., Lozovskaya, E. R. & Hartl, D. L. Molecular phylogeny and genome evolution in the *Drosophila virilis* species group: duplication of the alcohol dehydrogenase gene. *Mol. Biol. Evol.* **13**, 132–149 (1996).
- Patton, J. S., Gomes, X. V. & Geyer, P. K. Position-independent germline transformation in *Drosophila* using a cuticle pigmentation gene as a selectable marker. *Nucleic Acids Res.* **20**, 5859–5860 (1992).
- Karress, R. E. & Rubin, G. M. Analysis of P transposable element functions in *Drosophila*. *Cell* **38**, 135–146 (1984).
- Rubin, G. M. & Spradling, A. C. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353 (1982).

Acknowledgements. This work was supported by grants from the US Public Health Service. We thank W. Gilbert and S. R. Palumbi for their valuable suggestions and careful reading of the manuscript.

Correspondence and requests for materials should be addressed to D.I.N. (e-mail: dnurminsky@oeb.harvard.edu). All sequences have been deposited in GenBank under accession numbers AF070587–AF070689.

Ca²⁺/calmodulin signals the completion of docking and triggers a late step of vacuole fusion

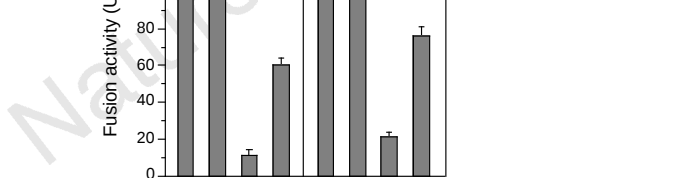
Christopher Peters & Andreas Mayer

Friedrich-Miescher-Laboratorium der Max-Planck-Gesellschaft, Spemannstrasse 37-39, 72076 Tübingen, Germany

The basic reaction mechanisms for membrane fusion in the trafficking of intracellular membranes and in exocytosis are probably identical¹. But in contrast to regulated exocytosis, intracellular fusion reactions are referred to as 'constitutive' as no final Ca²⁺-dependent triggering step has been observed. Although transport from the endoplasmic reticulum to the Golgi apparatus in the cell depends on Ca²⁺ (ref. 6), as does endosome fusion⁷ and assembly of the nuclear envelope⁸, it is unclear whether Ca²⁺ triggers these events. Membrane fusion involves several subreactions: priming, tethering and docking. Proteins that are needed for fusion include p115, SNAPS, NSF, SNAREs and small GTPases, which operate in these early reactions^{1–3}, but the machinery that catalyses the final mixing of biological membranes is still unknown. Here we show that Ca²⁺ is released from the vacuolar lumen following completion of the docking step. We have identified calmodulin as the putative Ca²⁺ sensor and as the first component required in the post-docking phase of vacuole fusion. Calmodulin binds tightly to vacuoles upon Ca²⁺ release. Unlike synaptotagmin or syncoilin in exocytosis⁴, calmodulin does not act as a fusion clamp but actively promotes bilayer mixing. Hence, activation of SNAREs is not sufficient to drive bilayer mixing between physiological membranes. We propose that Ca²⁺ control of the latest phase of membrane fusion may be a conserved feature, relevant not only for exocytosis, but also for intracellular, 'constitutive' fusion reactions. However, the origin of the Ca²⁺ signal, its receptor and its mode of processing differ.

Vacuole fusion can be reconstituted in a cell-free system⁹ and quantified by a biochemical complementation assay¹⁰. Three phases can be distinguished²: SNAP/NSF/LMA1-dependent priming leads to the ATP-driven disassembly of v/t-SNARE complexes^{11–13}, resulting in an activated state that is a prerequisite for docking, which is the second phase. Docking leads to the stable association of vacuoles and requires the Ras-like GTPase Ypt7p (ref. 13). No proteinaceous

The requirement for calmodulin is also apparent *in vivo*. Defects in components needed for fusion often result in abnormal vacuolar structure, such as fragmentation¹⁴ or enlargement of vacuoles³. We used a set of temperature-sensitive calmodulin mutants¹⁵ with point



mutations that affect the interaction with target proteins. These mutants fall into four intragenic complementation groups that affect distinct essential reactions¹⁵. This distinction may not hold for other, non-essential functions. We investigated strains from all four complementation groups for defects in vacuolar structure. Vacuoles were selectively stained *in vivo* with the fluorochrome FM4-64. The vacuoles of the strains DBY5708 (cmd1-226) and DBY5713 (cmd1-233) were slightly enlarged. DBY5708 (cmd1-228) showed few (<5%) cells with fragmented vacuoles (not shown). However, DBY5719 (cmd1-239) displayed a high proportion of cells with fragmented vacuoles just 2 h after the shift to the elevated temperature (34 °C; Fig. 2a). The vacuoles of wild-type cells remained unaltered under these conditions. The *in vivo* defect of this mutant strain was also reproduced *in vitro*. Isolated vacuoles from DBY5719 (grown at permissive temperature) were preincubated for 1 h at a restrictive temperature (37 °C). This treatment inactivated the vacuoles for fusion *in vitro* (Fig. 2b). Inactivated mutant vacuoles could be rescued *in vitro* by adding purified calmodulin. The rescued reaction was sensitive to microcystin LR, an inhibitor of a very late stage of vacuole fusion^{2,16}, indicating that the reaction still followed the authentic pathway. Comparable results were obtained if calmodulin was inactivated by temperature shift *in vivo*, that is, before vacuole isolation (not shown). Thus, the

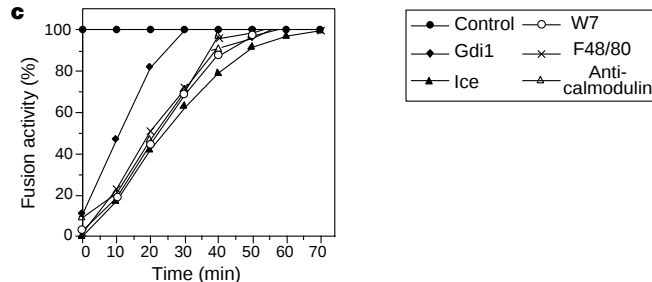


Figure 2 Calmodulin is required for a late stage of vacuole fusion *in vitro* and *in vivo*. **a**, *In vivo* phenotype. Yeast cells from strain DBY5719 (cmd1-239) or DBY5734 (CMD1, wild-type; WT) were grown in YPD at 25°C and stained with FM4-64. Then the cells were split, incubated at 25°C or 34°C (2 h) and analysed by confocal microscopy. The fluorescence and transmission channels were superimposed to visualize cell shape and the vacuoles. **b**, *In vitro* fusion. Vacuoles were prepared from DBY5719 (cmd1-239) and DBY5734 (CMD1) grown at permissive temperature (25°C). The vacuoles were preincubated (1 h, 37°C) in PS buffer with 150 mM KCl plus or minus purified calmodulin (10 µM). Then the samples were supplemented to yield fusion buffer and microcystin LR (8 µM) was added as indicated. After 1 h at 27°C fusion was assayed. **c**, Time course of calmodulin requirement. Six fusion samples were incubated at 27°C. At the indicated times they were divided into 30-µl standard reactions, supplemented with antagonists or buffer only, kept on ice for 10 min and then incubated at 27°C for the rest of the 70-min reaction period. One aliquot was kept on ice until the end of the reaction period to monitor the progression of fusion. The activity of the samples incubated at 27°C with addition of buffer only was set to 100%. The activity of the sample set on ice at 0 min is the 0% reference. Antagonist concentrations were: antibodies to calmodulin (5 µM), Gdi1 (2.5 µM), W7 (500 µM), F48/80 (500 µM).

fusion defect was a direct consequence of the inactivation of calmodulin.

Mutants of Ypt7 or Vam3, which are involved in docking, have fragmented vacuoles that are scattered throughout the cell^{14,17}. The fragments in DBY5719 (cmd1-239) are clustered together, consistent with a block in a post-docking event. We determined when calmodulin is required *in vitro* (Fig. 2c). Most vacuoles complete distinct reaction steps within defined intervals and become resistant to inhibitors of these steps^{2,16}. This approach can be used to investigate the sequence of events in vacuole fusion. Inhibitors or control buffer were added to an ongoing fusion reaction at different times and the samples were incubated further until the end of a standard fusion period (70 min). Another sample was set on ice to stop fusion at the given time and measure progression of the reaction. All inhibitors completely suppressed fusion when added at the start of the reaction (Fig. 2c). However, the reaction became resistant to Gdi1 within 30 min¹³. Gdi1 extracts the Ras-like GTPase Ypt7 from the vacuolar membrane¹⁷. Resistance of fusion to this protein indicates the completion of docking¹³. The inhibition curves for the antibodies specific to calmodulin and for W7 and F48/80 are close to the curve produced by placing the samples on ice (Fig. 2c). Therefore, calmodulin is required in a very late, post-docking phase of the reaction and acts close to the time of contents mixing.

Calmodulin bound to vacuoles in a Ca^{2+} -dependent manner (Fig. 3). A large proportion of the bound calmodulin was released from the vacuoles at Ca^{2+} concentrations below 500 nM, which includes the concentration range present in the cytosol of living yeast cells (100–150 nM)¹⁸. Released calmodulin could be recovered from the supernatant (not shown). Freshly isolated vacuoles (in the presence of 3 μM calcium) carry about 0.15 ng calmodulin per microgram vacuolar protein (estimated by western blotting; not shown), corresponding to 20–100 molecules per vacuole.

Some calmodulin functions do not depend on high-affinity binding of Ca^{2+} , for example, those in the spindle pole body¹⁹ or in endocytosis²⁰. Therefore, we used a calmodulin (cmd1-3) with point mutations inactivating two of the seven amino acids that coordinate calcium. This calmodulin has a greatly reduced affinity for Ca^{2+} *in vitro* ($K_d < 200\text{--}300\ \mu\text{M}$)²¹. Cells expressing it are temperature sensitive. Unlike wild-type calmodulin, the mutant protein could not reactivate vacuoles after a block imposed by the calmodulin antagonist W7 (Fig. 4a). If titrated into fusion samples of untreated vacuoles carrying wild-type calmodulin, the mutant calmodulin acted as a competitive inhibitor, with an IC_{50} of 25 μM (Fig. 4b). Wild-type calmodulin, when added at the same concentrations, was slightly stimulatory. Like wild-type calmodulin in the Ca^{2+} -free form, mutant calmodulin can still bind to its target on the vacuoles with low affinity. Because it is present in excess (a standard fusion sample contains only about 10 nM wild-type calmodulin), it can displace its wild-type counterpart. This indicates that it is in a functional conformation. *In vivo*, vacuoles of a strain carrying this mutant calmodulin were slightly enlarged at 25 °C. They displayed rapid fragmentation upon shift to 34 °C (not shown), in the same

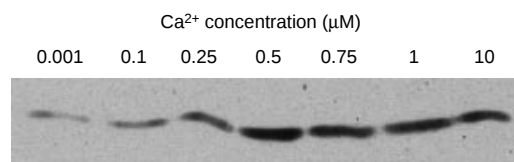


Figure 3 Calmodulin binds to vacuoles. Vacuoles (200 μl , 300 $\mu\text{g ml}^{-1}$) were suspended in fusion buffer (minus ATP and cytosol) with 1 mM BAPTA and CaCl_2 , yielding the indicated concentrations of free Ca^{2+} , and incubated (30 min, 27 °C). Vacuoles were pelleted (10,000g, 3 min, 4 °C), washed with the same buffer as before, pelleted and extracted with chloroform/methanol (1:1, 15 min, 25 °C). The samples were pelleted (20,000g, 10 min), dried (1 min, 95 °C), dissolved in sample buffer with 5 mM EGTA and analysed by SDS-PAGE and western blotting.

way as the cmd1-239 mutant (Fig. 2a). Vacuoles from the mutant strain (grown at 25 °C) were unable to fuse *in vitro* (Fig. 4c) but could be reactivated by adding purified wild-type calmodulin. Fusion of the reactivated vacuoles was sensitive to microcystin-LR, a very late-acting inhibitor of vacuole fusion, indicating that the rescued reaction still followed the authentic path. Thus, fusion incompetence was a direct consequence of the inability of calmodulin to bind Ca^{2+} and not due to indirect effects.

We investigated the role of Ca^{2+} by preincubating vacuoles with different inhibitors before using them in standard fusion assays (Fig. 5a). We used relatively high concentrations of ionomycin because ionomycin efficiently transports ions only above pH 7, but the vacuolar lumen is acidic. Moreover, free Mg^{2+} in the assay acts as a competitor. Ionomycin (a Ca^{2+} ionophore), thapsigargin and cyclopiazonic acid (inhibitors of Ca^{2+} -ATPases) strongly inhibited fusion. This indicates that a luminal Ca^{2+} pool must be actively maintained to keep the vacuoles fusion competent. Indeed, vacuoles accumulate Ca^{2+} by active transport to $>2\text{ mM}$ through an ATP-driven Ca^{2+} pump and a $\text{Ca}^{2+}/\text{H}^{+}$ antiporter¹⁸. BAPTA, which preferentially chelates Ca^{2+} , also suppressed fusion (Fig. 5b). In contrast, EGTA, which has a similar dissociation constant for Ca^{2+} but an on-rate much slower than BAPTA, was a very poor inhibitor. Inhibition was not due to Mg^{2+} depletion because BAPTA reduces the available Mg^{2+} concentration in the assay by less than 10%, which does not influence the reaction (data not shown). Moreover, the results were identical if Mg^{2+} was kept in molar excess over the chelator. BAPTA that had been presaturated with Ca^{2+} was not inhibitory. We also exclude the possibility of BAPTA interfering with the maturation or activity of alkaline phosphatase, the reporter enzyme of the fusion assay (see above and Fig. 6a). BAPTA, which is membrane impermeant, was only inhibitory if present throughout the reaction. Even after preincubating the vacuoles with 1 mM

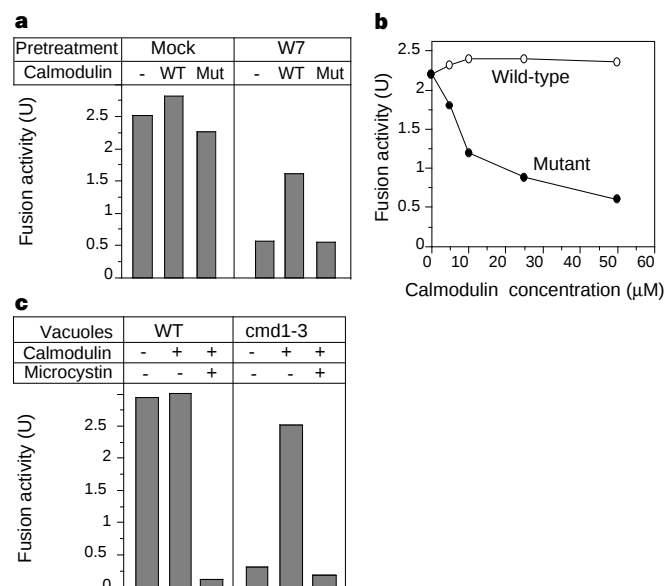


Figure 4 Calmodulin with a reduced affinity for Ca^{2+} inhibits fusion. **a**, Rescue from a block imposed by W7. Vacuoles were preincubated in PS/150 mM KCl plus or minus W7 (75 μM , 15 min, 0 °C), centrifuged (4,500g, 3 min, 2 °C) and resuspended in fusion buffer. They received control buffer only, 10 μM wild-type calmodulin, or 10 μM cmd1-3 calmodulin. After 70 min at 27 °C fusion was determined. **b**, Mutant calmodulin competitively inhibits fusion. Vacuoles were preincubated with wild-type or cmd1-3 calmodulin in fusion buffer minus ATP (60 min, 0 °C). Then the samples received ATP. Fusion was determined after 1 h at 27 °C. **c**, Vacuoles from the cmd1-3 mutant do not fuse. Standard fusion reactions were performed with vacuoles from strains JGY041(cmd1-13) and CRY1(CMD1, wt), grown at 25 °C, with wild-type calmodulin (10 μM) and microcystin LR (8 μM) as indicated.

BAPTA, the reaction was hardly influenced if the chelator was removed by reisolating the vacuoles before the fusion reaction was started (Fig. 5c). Together, these results indicate that the process relevant to fusion is efflux of Ca^{2+} from the vacuole lumen. Because of the high intra-vacuolar concentration of Ca^{2+} (2 mM), Ca^{2+} efflux could lead to drastic transient increases in the Ca^{2+} concentration close to the vacuolar surface, which require a fast chelator in sufficiently high concentrations to be quenched rapidly. Furthermore, this is consistent with a close spatial arrangement of the calcium release site and the site of calcium action.

Can the vacuoles be rescued from a block imposed by the withdrawal of luminal Ca^{2+} ? We preincubated vacuoles with BAPTA in the presence of ionomycin in order to deplete luminal Ca^{2+} (Fig. 5d). The vacuoles were re-isolated and resuspended in fresh fusion buffer without BAPTA. In this buffer the pretreated vacuoles were unable to fuse. Control vacuoles pretreated with BAPTA or buffer only (not depleted of luminal Ca^{2+}) were fusion competent. Fusion competence of the Ca^{2+} -depleted vacuoles was restored if the buffer was supplemented with Ca^{2+} . Ca^{2+} added after re-isolation is rapidly taken up by the vacuoles at the start of the fusion reaction (compare Fig. 6b) and restored the depleted luminal pool (not shown). The vacuolar lumen is about 5% of the total liquid volume in our *in vitro* reaction (unpublished observation). Therefore, Ca^{2+} -depleted vacuoles need a Ca^{2+} concentration of 100 μM in the medium to reconstitute their normal luminal Ca^{2+} concentration of >2 mM by active uptake. Addition of higher

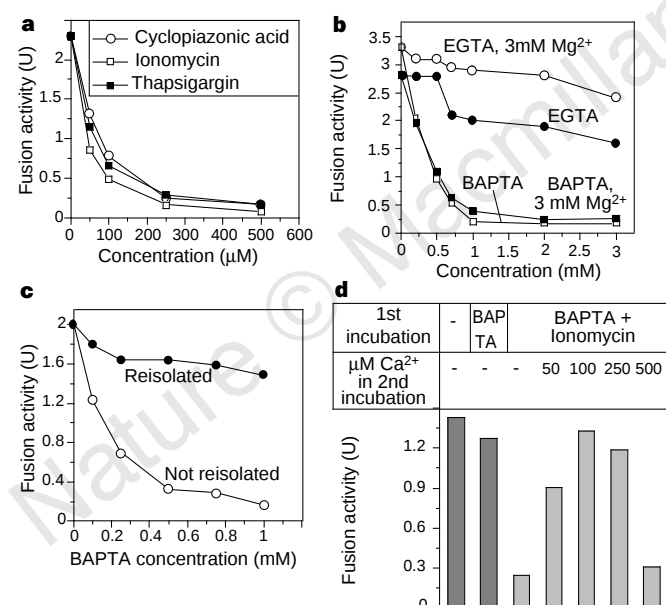


Figure 5 Luminal Ca^{2+} is required for vacuole fusion. **a**, Influence of ionomycin and calcium ATPase inhibitors. Vacuoles were preincubated (15 min, 0°C) in fusion buffer without ATP with the indicated concentrations of cyclopiazonic acid, ionomycin or thapsigargin. Then ATP was added to start fusion. After 70 min at 27°C, fusion was assayed. **b**, Influence of Ca^{2+} chelators. Vacuoles were incubated in fusion buffers (minus ATP) at the given chelator concentrations (5 min, 0°C) plus or minus 3 mM MgCl_2 . Then ATP was added and after 70 min at 27°C fusion was assayed. **c**, BAPTA inhibits only if present throughout the reaction. Samples were preincubated with BAPTA (at standard Mg^{2+} concentration) as in **b**. Open circles, the fusion reaction was started directly by adding ATP. Closed circles, the vacuoles were centrifuged (4,500g, 3 min, 4°C) and resuspended in fresh fusion buffer without BAPTA, but with ATP. After 70 min at 27°C, fusion was assayed. **d**, Rescue of fusion after depleting luminal Ca^{2+} . Vacuoles were incubated in fusion buffer without cytosol in the presence of BAPTA (5 mM), BAPTA (5 mM) plus ionomycin (100 μM), or control buffer only (30 min, 27°C). After centrifugation (4,500g, 3 min, 4°C), the vacuoles were resuspended in fusion buffer plus 10 μM calmodulin and supplemented with CaCl_2 as indicated. Fusion was assayed after 70 min at 27°C.

calcium concentrations exceeds the vacuolar uptake capacity and inhibits fusion (not shown).

We determined at which stage Ca^{2+} acts by asking when the reaction would become resistant to appropriate inhibitors. Fusion was sensitive to ionomycin and BAPTA throughout the incubation, much longer than to the docking marker Gdi1 (Fig. 6a). Therefore, Ca^{2+} is required in the post-docking phase of vacuole fusion, close to contents mixing. We also measured the extra-vacuolar concentration of Ca^{2+} in the course of the reaction by pelleting the vacuoles quickly and assaying the supernatant for Ca^{2+} (Fig. 6b). Within the first 5–10 min the Ca^{2+} concentration in the solution dropped dramatically. This drop was ATP dependent and reflects active uptake of Ca^{2+} through the vacuolar Ca^{2+} pump and $\text{Ca}^{2+}/\text{H}^{+}$ exchange protein¹⁸. Later the Ca^{2+} concentration in the supernatant increased transiently, demonstrating net Ca^{2+} release from the vacuoles. Net Ca^{2+} efflux, which becomes detectable only after excess Ca^{2+} in the reaction mixture has been taken up by the vacuoles, depends on priming and docking: it did not occur in the presence of Gdi1 (preventing docking) or antibodies to Sec18 (preventing fusion)¹³. In the late phase (50 min) calcium uptake prevailed again because no new docking events are initiated and calcium release is no longer triggered. The slight increase of the Ca^{2+} concentration at the last time point is due to lysis of a small fraction of vacuoles that can occur after longer incubations at 27°C (A.M., unpublished observation). Net release of Ca^{2+} and the time course of fusion do not directly correspond to each other. This was

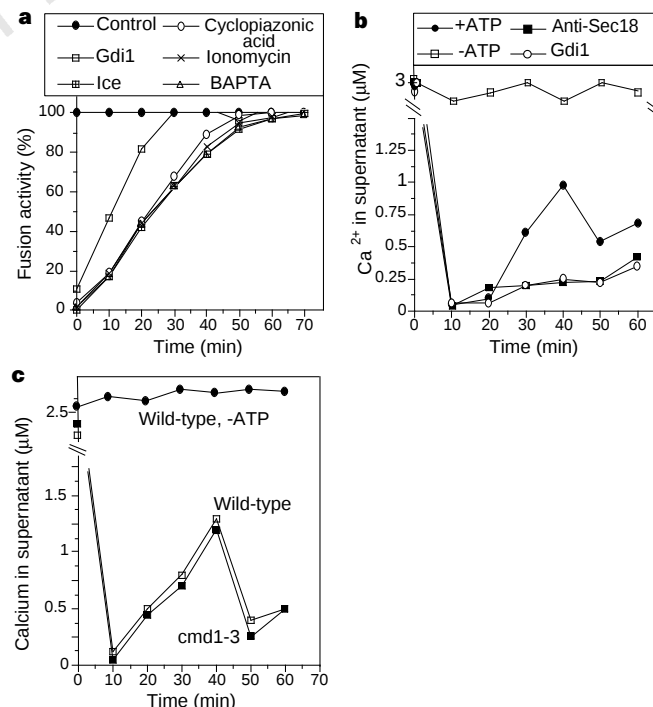


Figure 6 The efflux of Ca^{2+} from the lumen is needed in a late phase of fusion. **a**, Time course of Ca^{2+} requirement. The experiment was performed as in Fig. 2c. Inhibitors were added into ongoing fusion reactions at different times after the start and the influence of the inhibitors on the rest of the incubation period was monitored. **b**, Ca^{2+} concentration in the medium. Vacuoles (30× a standard fusion sample) were incubated in fusion buffer minus ATP with Gdi1 or antibodies to Sec18, or with control buffer (15 min, 0°C). The reaction was started by ATP and transferred to 27°C. A control sample was incubated without ATP. At different times 100- μl aliquots were removed and centrifuged (7,700g, 3 min, 4°C). Ca^{2+} in the supernatant was determined. Inhibitor concentrations: Gdi1 (2.5 μM), antibodies to Sec18 (1.5 μM), BAPTA (3 mM), cyclopiazonic acid (500 μM), thapsigargin (500 μM), ionomycin (500 μM). **c**, Ca^{2+} uptake and efflux in mutant vacuoles. Fusion reactions with vacuoles from JGY041 (cmd1-3) and CRY1 (WT) were analysed for Ca^{2+} in the supernatant as in **b**.

expected for the following reasons. First, excess of Ca^{2+} in the medium, caused by slow leakage from the vacuoles during preparation, prevents detection of Ca^{2+} efflux during the first 10 min. Second, events on individual vacuoles are obscured by the behaviour of the whole population. Third, the fusion assay cannot detect repeated rounds of fusion, whereas the determination of Ca^{2+} in the supernatant does. We conclude that the efflux of Ca^{2+} from the vacuoles is controlled by the preceding steps of the overall reaction, priming and docking.

Could calmodulin act as a receptor for this Ca^{2+} signal? Vacuoles from the *cmd1-3* mutant were indistinguishable from wild-type organelles in their uptake and release of Ca^{2+} (Fig. 6c). However, the *cmd1-3* protein does not support fusion. Therefore, the defect of this mutant calmodulin in Ca^{2+} binding must affect the reception of the Ca^{2+} signal rather than its generation. This indicates that calmodulin is required as the receptor for the Ca^{2+} signal which triggers bilayer fusion.

We have integrated our new findings with previous results into the following model. In the course of priming, pre-existing v/t-SNARE complexes are split by Sec17 (α -SNAP) and Sec18 (NSF). This activated state is stabilized by LMA1. It leads to docking and the reformation of SNARE complexes. Ypt7 is required in this process until docking is completed. We postulate that docking produces an unknown signal leading to the transient opening of Ca^{2+} channels. The resulting Ca^{2+} efflux is perceived by calmodulin, which then binds to the vacuoles more tightly and activates enzymes that catalyse bilayer/contents mixing. Because calmodulin is not enriched on vacuoles of a living yeast cell²², this Ca^{2+} -dependent increase in affinity may serve to recruit it to the membrane in response to the Ca^{2+} signal. Unlike the 'fusion clamps' synaptotagmin or syncoilin⁴, which interfere with fusion at low Ca^{2+} concentration, calmodulin exerts an active, promoting effect on the post-docking events. Therefore, SNARE priming by NSF and SNAP cannot be sufficient to fuse physiological membranes. In contrast, artificial proteoliposomes containing purified SNAREs can fuse with a low rate and efficiency in the absence of other factors²³.

Although the Ca^{2+} signal appears to fulfil the same terminal function as it does in exocytosis, its generation is controlled by the reaction sequence itself, rather than by extracellular stimuli. Coordinating docking with the late reaction of bilayer/contents mixing could be crucial. The reaction sequence may involve labile, activated intermediates that must be used rapidly to be productive. Such a labile state is also induced during priming^{12,13}. Lability could serve as a safeguard against promiscuous fusion events which might result, for example, from accidental 'undocking' of activated fusion partners before bilayer fusion has been achieved. Such undocking events have been observed microscopically²⁴ and may correspond to the decay of the primed/docked state of secretory vesicles in the absence of ATP²⁵.

Our finding is consistent with earlier reports linking calmodulin to regulated exocytosis²⁶, transcytosis²⁷ and endocytotic membrane trafficking^{7,20}. There, however, it was not resolved if calmodulin was required in an early or late phase of fusion, or whether it affected membrane trafficking by other means. For regulated secretion from chromaffin cells, the calmodulin requirement was mapped to the triggering stage²⁸, but this finding is controversial²⁹. In vacuole-fusion, calmodulin clearly has a function after the docking event. It is the first protein of the post-docking phase to be identified and it could be a key to other late-acting factors catalysing bilayer/contents mixing. □

Methods

Yeast strains and chemicals. The following strains were used: BJ3505 and DKY6281¹⁰; DBY5734/YOC200 (*CMD1*)¹⁵; DBY5706/YOC226 (*cmd1-226*), DBY5708/YOC228 (*cmd1-228*), DBY5713/YOC233 (*cmd1-233*) and DBY5719/YOC239 (*cmd1-239*); CRY1 (*CMD1*)¹⁹ and JGY41 (*cmd1-3*). *Pho8* was deleted using plasmid pAR2 (J. Shaw, Salt Lake City, Utah) and *Pep4* using

plasmid pTS17 or pTS15 (T. Stevens, Eugene, Oregon). Deletions were verified by western blotting and PCR. Chemicals: W5, W7, F48/80, ophiobolin A, thapsigargin, cyclopiazonic acid, ionomycin, BAPTA (Calbiochem); EGTA (Roth, Karlsruhe); water with a low Ca^{2+} content (Merck); for others, see ref. 2.

Biochemical procedures. Gdi1 from *Escherichia coli* BL21 (pNB620) and affinity-purified IgG were made as described². Anti-calmodulin IgG was eluted rapidly and neutralized immediately. Yeast calmodulin was linked to activated CH-Sepharose 4B (Pharmacia) as follows: 5 mg purified calmodulin was mixed with 1 ml Sepharose in 0.1 M NaHCO_3 , pH 8, 0.15 M NaCl, rotated (4 h, 4 °C), blocked with 1 M Tris-Cl, pH 8 (1 h, 22 °C), washed three times with PBS and stored in PBS/1 mM sodium azide. Stock solutions were: W5 (5 mM in PS buffer: 10 mM PIPES/KOH, pH 6.8, 200 mM sorbitol), W7 (25 mM in DMSO), F48/80 (5 mM in DMSO), ophiobolin (5 mM in methanol), thapsigargin, cyclopiazonic acid, ionomycin (10 mM in DMSO), BAPTA, EGTA (100 mM in PS), microcystin-LR (2 mM in DMSO) and FM4-64 (10 mM in DMSO). All stocks were kept at -20 °C.

Vacuole fusion. Vacuole² and cytosol¹³ preparation and fusion² (MnCl_2 was omitted) followed published procedures. Alkaline phosphatase activity was assayed as described¹⁰. However, 10 mM CaCl_2 was added to the developing solution to avoid interference of chelators with alkaline phosphatase activity. The value of a control sample kept at 0 °C (which prevents fusion) was subtracted from all samples. Fusion activity of 1 U is defined as 1 μmol *p*-nitrophenol developed per min and μg *Dpep4* vacuoles at 27 °C.

Determination of Ca^{2+} in the fusion assay. A sample (100 μl) of a fusion reaction was centrifuged (7,700g, 3 min, 4 °C); 50 μl supernatant was supplemented with 10–25 μM BAPTA and 50 nM aequorin and immediately measured in a luminometer (Berthold LB 9501). The Ca^{2+} concentration of the sample was determined by calibration curves obtained by adding defined concentrations of CaCl_2 to fusion buffer. This fusion buffer had been depleted of Ca^{2+} before use by incubation with 0.2 mg ml^{-1} vacuoles (10 min, 27 °C), followed by pelleting the membranes (10 min, 14,000 g).

Purification of calmodulin. Wild-type calmodulin (pJG028) and *cmd1-3* calmodulin (pJG031) were expressed in *E. coli* GM1²¹. Wild-type calmodulin was purified by phenylsepharose chromatography²². For experiments where *cmd1-3* and wild-type calmodulin were compared, both proteins were purified by native PAGE (in a standard 15% gel without SDS) in a Bio-Rad Prep Cell. Eluted proteins were transferred into PS with 150 mM KCl by repeated dilution and reconcentration in a Centrprep-3 ultrafiltration device. They were stored at -20 °C at a concentration of 8 mg ml^{-1} .

Microscopic assay of yeast cells. Cells were shaken in 5 ml YPD overnight (20 ml tubes, 30 °C, 225 r.p.m.). They were diluted to an absorbance (*A*) at 600 nm of 0.1 with YPD, grown for 6 h and diluted to *A*₆₀₀ = 0.4 with YPD. A sample (1 ml) of the suspension was supplemented with 20 μM FM4-64³⁰, transferred to new tubes and incubated as above for 1 h. The cells were collected (2 min, 1,300g), washed with 1 ml YPD, pelleted as above, resuspended in 1 ml YPD and shaken for another 30 min. When samples were drawn, the cells were immediately chilled on ice, reisolated (2 min, 1,300g, 4 °C) and resuspended in YPD at an *A*₆₀₀ of 3–4. A sample (7 μl) of the suspension was mixed with 7 μl of 0.4% Seaplaque agarose in PS (kept liquid at 35 °C), transferred to a slide and chilled at 4 °C for 5 min to immobilize the cells. Cells were quickly investigated with a confocal microscope (Leica TCS) under minimal excitation.

Received 10 August; accepted 9 October 1998.

1. Cao, X., Ballew, N. & Barlowe, C. Initial docking of ER-derived vesicles requires Uso1p and Ypt1p but is independent of SNARE proteins. *EMBO J.* **17**, 2156–2165 (1998).
2. Mayer, A., Wickner, W. & Haas, A. Sec18p (NSF) driven release of Sec17p (a-SNAP) can precede docking and fusion of yeast vacuoles. *Cell* **85**, 83–94 (1996).
3. Nichols, B. J., Ungermann, C., Pelham, H. R. B., Wickner, W. T. & Haas, A. Homotypic vacuolar fusion mediated by t- and v-SNAREs. *Nature* **387**, 199–202 (1997).
4. Edwardson, J. M., An, S. & Jahn, R. The secretory granule protein syncoilin binds to syntaxin in a Ca^{2+} -sensitive manner. *Cell* **90**, 325–333 (1997).
5. Rothman, J. E. Mechanisms of intracellular protein transport. *Nature* **372**, 55–63 (1994).
6. Pryer, N. K., Wuestehube, L. J. & Schekman, R. Vesicle-mediated protein sorting. *Annu. Rev. Biochem.* **61**, 471–516 (1992).
7. Colombo, M. I., Beron, W. & Stahl, P. D. Calmodulin regulates endosome fusion. *J. Biol. Chem.* **272**, 7707–7712 (1997).
8. Sullivan, K. M., Busa, W. B. & Wilson, K. L. Calcium mobilization is required for nuclear vesicle fusion in vitro: implications for membrane traffic and IP₃ receptor function. *Cell* **73**, 1411–1422 (1993).
9. Conradt, B., Shaw, J., Vida, T., Emr, S. & Wickner, W. In vitro reactions of vacuole inheritance in *Saccharomyces cerevisiae*. *J. Cell Biol.* **119**, 1469–1479 (1992).
10. Haas, A., Conradt, B. & Wickner, W. G-protein ligands inhibit in vitro reactions of vacuole inheritance. *J. Cell Biol.* **126**, 87–97 (1994).

11. Ungermann, C., Nichols, B. J., Pelham, H. R. B. & Wickner, W. A vacuolar v-t-SNARE complex, the predominant form in vivo and on isolated vacuoles, is disassembled and activated for docking and fusion. *J. Cell Biol.* **140**, 61–69 (1998).
12. Xu, Z., Mayer, A., Muller, E. & Wickner, W. A heterodimer of thioredoxin and I^{β} cooperates with Sec18p (NSF) to promote yeast vacuole inheritance. *J. Cell Biol.* **136**, 299–306 (1997).
13. Mayer, A. & Wickner, W. Docking of yeast vacuoles is catalyzed by the rad-like GTPase Ypt7p after symmetric priming by Sec18p (NSF). *J. Cell Biol.* **136**, 307–317 (1997).
14. Wada, Y., Ohsumi, Y. & Anraku, Y. Genes for directing vacuolar morphogenesis in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **267**, 18665–18670 (1992).
15. Ohya, Y. & Botstein, D. Diverse essential functions revealed by complementing yeast calmodulin mutants. *Science* **263**, 963–966 (1994).
16. Conradt, B., Haas, A. & Wickner, W. Determination of four biochemically distinct, sequential stages during vacuole inheritance in vitro. *J. Cell Biol.* **126**, 99–110 (1994).
17. Haas, A., Schlegelmann, D., Lazar, T., Gallwitz, D. & Wickner, W. The GTPase Ypt7 of *Saccharomyces cerevisiae* is required on both partner vacuoles for the homotypic fusion step of vacuole inheritance. *EMBO J.* **14**, 5258–5270 (1995).
18. Cunningham, K. W. & Fink, G. R. Ca^{2+} transport in *Saccharomyces cerevisiae*. *J. Exp. Biol.* **196**, 157–166 (1994).
19. Geiser, J. R., Sundberg, H. A., Chang, B. H., Muller, E. G. D. & Davis, T. N. The essential mitotic target of calmodulin is the 110-kilodalton component of the spindle pole body in *S. cerevisiae*. *Mol. Cell. Biol.* **13**, 7913–7924 (1993).
20. Kibbler, E., Schimmöller, F. & Riezmann, H. Ca^{2+} -independent calmodulin requirement for endocytosis in yeast. *EMBO J.* **13**, 5539–5546 (1994).
21. Geiser, J. R., van Tuinen, D., Brockerhoff, S. E., Neff, M. & Davis, T. N. Can calmodulin function without Ca^{2+} ? *Cell* **65**, 949–959 (1991).
22. Brockerhoff, S. E. & Davis, T. N. Calmodulin concentrates at regions of cell growth in *Saccharomyces cerevisiae*. *J. Cell Biol.* **118**, 619–629 (1992).
23. Weber, T. et al. SNAREpins: minimal machinery for membrane fusion. *Cell* **92**, 759–772 (1998).
24. Steyer, J. A., Horstmann, H. & Almers, W. Transport, docking and exocytosis of single secretory granules in live chromaffin cells. *Nature* **388**, 474–478 (1997).
25. Holz, R. W., Bittner, M. A., Peppers, S. C., Senter, R. A. & Eberhardt, D. A. MgATP-independent and MgATP-dependent exocytosis. *J. Biol. Chem.* **264**, 5412–5419 (1989).
26. Steinhardt, R. A. & Alderton, J. M. Calmodulin confers Ca^{2+} sensitivity on secretory exocytosis. *Nature* **295**, 154–155 (1982).
27. Apodaca, G., Enrich, C. & Mostov, K. E. The calmodulin antagonist, W-13, alters transcytosis, recycling, and the morphology of the endocytic pathway in MDCK cells. *J. Biol. Chem.* **269**, 19005–19013 (1994).
28. Kibble, V. A. & Burgoyne, R. D. Calmodulin increases the initial rate of exocytosis in adrenal chromaffin cells. *Eur. J. Physiol.* **431**, 464–466 (1996).
29. Artalejo, C. R., Elhamdani, A. & Palfrey, H. C. Calmodulin is the divalent cation receptor for rapid endocytosis, but not exocytosis, in adrenal chromaffin cells. *Neuron* **16**, 195–205 (1996).
30. Vida, T. & Emr, S. D. A new vital stain for visualizing vacuolar membrane dynamics and endocytosis in yeast. *J. Cell Biol.* **128**, 779–792 (1995).

Acknowledgements. We thank T. Davis for many strains and stimulating discussions; B. Wickner for the start-up kit of reagents; D. Botstein for the calmodulin ts mutants; and P. Overath and U. Henning for their support and the opportunity to use their facilities. We thank O. Müller, P. Overath and B. Wickner for critically reading the manuscript and A. Glatz and C. Baradon for assistance. This work was supported by the Boehringer Ingelheim Foundation and the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to A.M. (e-mail: Andreas.Mayer@Tuebingen.mpg.de).

Control of apoptosis and mitotic spindle checkpoint by survivin

Fengzhi Li[†], Grazia Ambrosini[†], Emily Y. Chu^{*}, Janet Plescia^{*}, Simona Tognin[‡], Pier Carlo Marchisio[‡] & Dario C. Altieri^{*}

^{*} Boyer Center for Molecular Medicine, Department of Pathology, Yale University School of Medicine, 295 Congress Avenue, New Haven, Connecticut 06536, USA
[‡] DIBIT S. Raffaele Scientific Institute, Via Olgettina 58, Milano 20132, Italy
[§] Department of Biomedical Sciences and Human Oncology, University of Torino, Torino 10126, Italy

[†] These authors contributed equally to this work

Progression of the cell cycle and control of apoptosis (programmed cell death) are thought to be intimately linked processes¹, acting to preserve homeostasis and developmental morphogenesis². Although proteins that regulate apoptosis have been implicated in restraining cell-cycle entry³ and controlling ploidy (chromosome number)⁴, the effector molecules at the interface between cell proliferation and cell survival have remained elusive. Here we show that a new inhibitor of apoptosis (IAP) protein^{5,6}, survivin⁷, is expressed in the G2/M phase of the cell cycle in a cycle-regulated manner. At the beginning of mitosis, survivin associates with microtubules of the mitotic spindle in a specific and saturable reaction that is regulated by microtubule

dynamics⁸. Disruption of survivin–microtubule interactions results in loss of survivin’s anti-apoptosis function and increased caspase-3 activity, a mechanism involved in cell death, during mitosis. These results indicate that survivin may counteract a default induction of apoptosis in G2/M phase. The overexpression of survivin in cancer⁷ may overcome this apoptotic checkpoint and favour aberrant progression of transformed cells through mitosis.

Undetectable in quiescent tissues, survivin was abundantly expressed in proliferating cells and rapidly downregulated by cell-cycle arrest in G1 phase (data not shown). We therefore investigated the potential cell-cycle dependence of survivin expression. Treatment with mimosine, thymidine and nocodazole arrested HeLa cells in late G1 (80%), S (64%) and G2/M (66%), respectively, as determined by DNA-content analysis (data not shown). Endogenous survivin RNA was undetectable in HeLa cells in G1, increased 6.2-fold in S-phase cells, and was upregulated ~40-fold in G2/M cells,

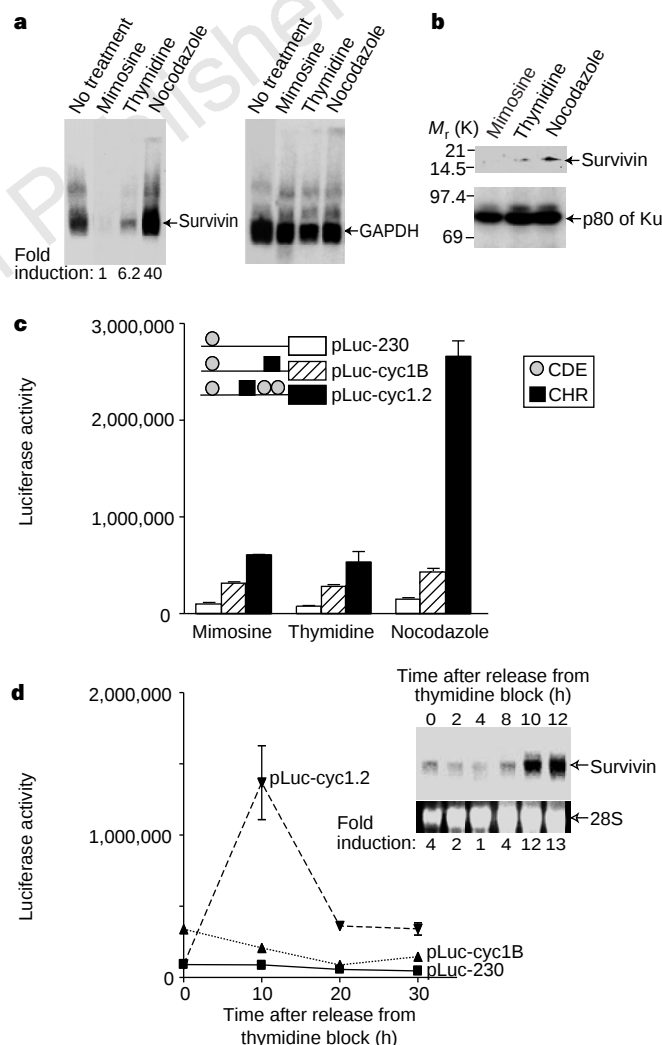


Figure 1 Cell-cycle-dependent expression of survivin in G2/M. **a**, Northern blot of survivin (left) or GAPDH (right) RNA in exponentially growing (no treatment) or drug-synchronized HeLa cells (mimosine, G1; thymidine, S; nocodazole, G2/M). **b**, Immunoblotting of survivin or of the p80 subunit of Ku protein in drug-synchronized HeLa cells. **c**, Survivin promoter activity in drug-synchronized HeLa cells. Inset, map of survivin-luciferase promoter constructs containing various CDE/CHR transcriptional regulatory elements⁹. **d**, Time course of survivin promoter activity in HeLa cells after release from a thymidine-induced G1/S block. Inset, Northern blot of endogenous survivin RNA in cell-cycle-synchronized HeLa cells. Data are the mean $\pm$ s.d. of duplicates of a representative experiment out of at least three independent determinations. 28S, 28S RNA.

whereas expression of GAPDH RNA was unchanged (Fig. 1a). Survivin protein (relative molecular mass 16,500 (M_r 16.5K)) but not a control protein, the p80 subunit of the Ku antigen, was also shown by immunoblotting to be upregulated in G2/M (Fig. 1b). Inspection of the survivin promoter⁷ revealed the presence of three potential CDE (sequence GGCGG at positions -6, -12 and -171) and one CHR (sequence ATTTGAA at position -42) G1 transcriptional repressor elements, which control cell-cycle periodicity of expression of various G2/M-regulated genes⁹. A pLuc-cycl2 construct, containing all four CDE and CHR elements upstream of a luciferase reporter gene, directed a roughly five- to tenfold increase in promoter activity in G2/M-synchronized HeLa cells, as compared with cells synchronized in G1 and S phase (Fig. 1c). In contrast, deletion of the proximal CDE elements in pLuc-230 and pLuc-cycl1B produced minimal luciferase activity without cell-cycle dependence (Fig. 1c). Maximal survivin promoter activity and peak levels of endogenous survivin RNA in HeLa cells coincided with G2/M phase, 10 hours after release from thymidine-induced arrest in G1/S phase (Fig. 1d and see below).

We investigated whether there might be a relationship between survivin expression in G2/M and assembly of the mitotic apparatus. We studied cells from prophase onwards; the anti-survivin monoclonal antibody 8E2 or the affinity-purified antibody to the survivin sequence A3-I19 (ref. 7) indistinguishably decorated Taxol-stabilized spindle fibres at metaphase (Fig. 2a), anaphase (Fig. 2b) and late telophase in midbodies (Fig. 2d) in methanol/acetone-fixed cells. Similar results were obtained in the absence of Taxol, although spindle fibres were blurred and less well identified. This pattern of survivin distribution was identical to the pattern seen with the anti- β -tubulin monoclonal antibody 20C6, with simultaneous chromosomal staining at anaphase (Fig. 2c) and at telophase during cytodieresis (Fig. 2e). Disruption of microtubules by colchicine resulted in dispersion of survivin immunoreactivity among metaphase chromosomes (Fig. 2f). In interphase cells, survivin was associated with γ -tubulin around spindle centrioles (data not shown) and, upon stabilization of GTP, some survivin decorated cytoplasmic microtubules radiating from microtubule-organizing centres (MTOCs; see below and Figs 3f, 4f).

We characterized the survivin-tubulin interaction further: in co-sedimentation experiments, a survivin band of M_r 16.5K was immunoblotted in pellets of Taxol-stabilized microtubules from HeLa cell extracts or in brain-purified tubulin (Fig. 2g, lanes 1, 2), in association with Coomassie-blue-stained tubulin (M_r 55K) (Fig. 2g, lanes 5, 6). In contrast, microtubules depolymerized with nocodazole did not contain survivin (Fig. 2g, lanes 3, 4) or tubulin (Fig. 2g, lanes 7, 8). In control experiments, p80 of Ku did not associate with Taxol-polymerized microtubules (Fig. 2h). Concentration-dependent binding of survivin to polymerized tubulin approached saturation at $\sim 1.8 \mu\text{M}$ of added survivin (Fig. 2i), with an affinity (K_d) of $\sim 5-7 \mu\text{M}$ (Fig. 2j) which is comparable with the affinity of conventional microtubule-associated proteins, such as tau¹⁰, for polymerized tubulin. The survivin-tubulin interaction also depended on microtubule dynamics⁸, with reversible enrichment of survivin in pellets or supernatants during successive rounds of microtubule polymerization and depolymerization, respectively (Fig. 2k).

We next determined whether survivin-microtubule interactions were required for apoptosis inhibition. Transfection of wild-type survivin or the apoptosis inhibitors X-linked IAP (XIAP)¹¹ or Bcl-2 (ref. 12) protected NIH3T3 fibroblasts from apoptosis induced by microtubule-stabilizing Taxol (Fig. 3a). In contrast, a truncated survivin coil-less mutant consisting of residues M1-G99, which lacks the charged carboxy-terminal coiled-coil⁷, was not cytoprotective (Fig. 3a). In co-sedimentation experiments, the survivin coil-less mutant bound minimally (8% of wild-type) to polymerized microtubules, accumulating in supernatants of Taxol-treated samples (Fig. 3b). Accordingly the survivin coil-less mutant promi-

nently accumulated in the cytoplasm of NIH3T3 transfectants, but failed to associate with mitotic spindle microtubules, as shown by immunofluorescence (Fig. 3c). In contrast, wild-type survivin bound to mitotic spindle microtubules during metaphase (Fig. 3d) and late telophase in midbodies (Fig. 3e), and was associated with

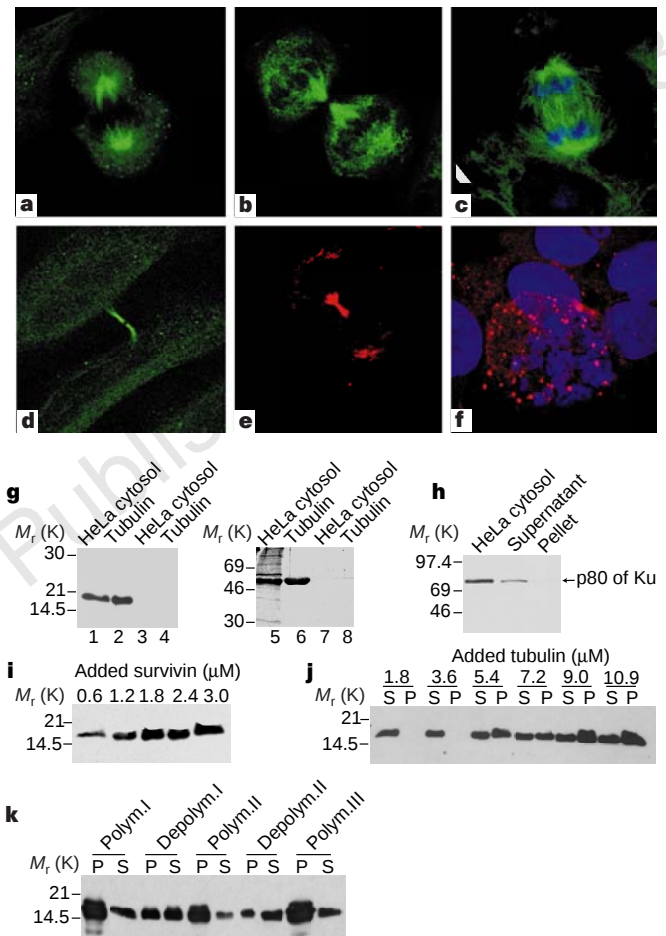


Figure 2 Survivin-microtubule interaction. HeLa cells were stained with the anti-survivin monoclonal antibody 8E2 (ref. 21) (**a, b, d, f**) or anti- β -tubulin monoclonal antibody 20C6 (**c, e**) after Taxol (**a-e**) or colchicine (**f**) treatment. **c**, Simultaneous staining of β -tubulin (monoclonal antibody 20C6, green) and anaphase chromosomes (Hoechst 33342, blue). Immunodetection of tubulin and survivin in **e** and **f**, respectively, was carried out with a Texas-red-conjugated second antibody. For other panels, a fluorescein isothiocyanate (FITC)-conjugated second antibody (green) was used. **a, f**, Metaphase (Hoechst 33342 stains chromosomes blue in **f**); **b, c**, anaphase; **d, e**, midbodies in late telophase. Pictures were taken with a BioRad 1024 confocal microscope (**a, b, d, e**) or with a Zeiss Axiophot microscope (**c, f**), all with a $\times 63$ lens. **g**, Survivin (M_r 16.5K) is shown by immunoblot to be found in co-sedimented pellets of Taxol-stabilized polymerized microtubules from HeLa cells (lane 1) or brain-purified tubulin (lane 2), but does not localize with nocodazole-depolymerized microtubules (lanes 3, 4). Coomassie-blue-stained tubulin (M_r ~ 55 K) co-associates with survivin in Taxol-treated pellets (lanes 5, 6), but not in nocodazole-depolymerized microtubules (lanes 7, 8). **h**, Enrichment of the p80 subunit of Ku in the supernatant but not in the pellet (which contains microtubules) of Taxol-stabilized HeLa cell microtubules. **i**, Concentration dependence and saturability ($1.8 \mu\text{M}$) of survivin binding to polymerized microtubules. **j**, Affinity of survivin for polymerized microtubules at increasing tubulin concentrations. A K_d of $\sim 5-7 \mu\text{M}$ was determined (the tubulin concentration required to bind 50% of total survivin added¹⁰). **k**, Regulation of survivin-microtubule interactions by microtubule dynamics⁸ during successive rounds of temperature-dependent microtubule polymerization (Polym.) or depolymerization (Depolym.). For all co-sedimentation experiments in **g-k**, samples in pellets (P) or supernatants (S) were immunoblotted with the affinity-purified antibody to the survivin sequence A3-I19 (ref. 7).

the MTOC in interphase cells transfected with survivin (Fig. 3f, arrow). Next, we compared the effects of Bcl-2 or survivin on apoptosis induced by microtubule-interfering drugs. Survivin efficiently inhibited Taxol-induced apoptosis, but was ineffective against microtubule-depolymerizing nocodazole or vincristine (Fig. 3g). In contrast, Bcl-2 inhibited apoptosis induced by all three drugs, independently of microtubule assembly/disassembly (Fig. 3g). Lack of survivin-mediated cytoprotection was observed at

the lowest concentration of vincristine, whereas Bcl-2 was anti-apoptotic over a wide range of vincristine concentrations (Fig. 3g, inset).

We next focused on the survivin baculovirus *iap* repeat (BIR)⁷, which in other IAP proteins is required for inactivation of terminal caspase-3 and caspase-7 (refs 11, 13). Mutation of Cys 84 → Ala (C84A) in the survivin BIR resulted in complete loss of cytoprotection against Taxol-induced apoptosis, whereas Pro 26 → Ala or

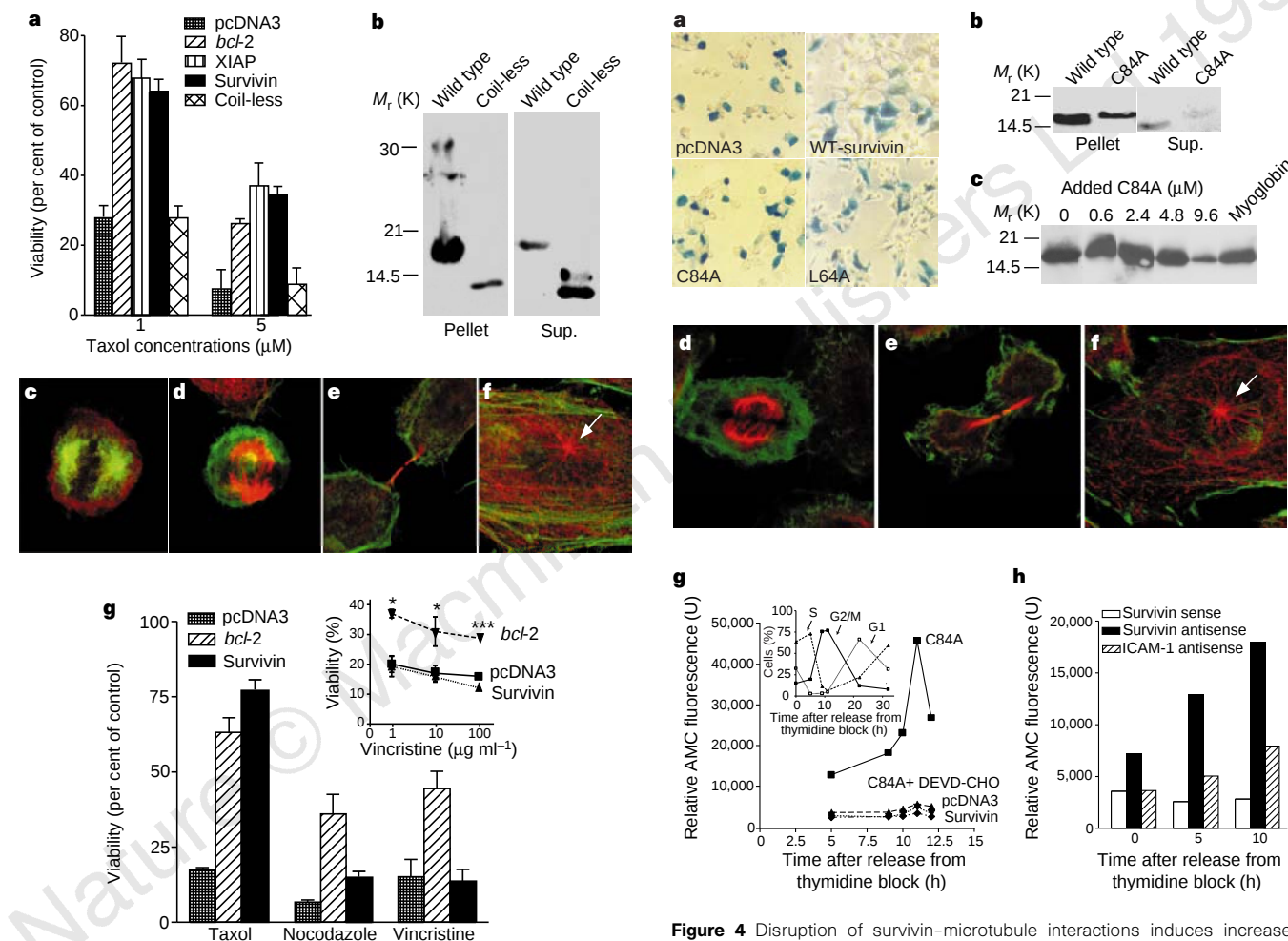


Figure 3 Survivin-microtubule interaction is required for apoptosis inhibition. **a**, Taxol-induced apoptosis in NIH3T3 transfectants is inhibited by wild-type survivin, Bcl-2 and XIAP, but not by the survivin coil-less mutant M1-G99. Wild-type and mutant survivin were expressed at comparable levels as shown by immunoblotting. pcDNA3 is a control vector. **b**, Reduced binding of recombinant coil-less survivin to polymerized microtubules (pellet; 8% of wild-type binding) and relative enrichment in supernatants of Taxol-treated samples (Sup.). **c**, Immunolocalization of survivin coil-less mutant (red) in the cytoplasm of MSB-fixed NIH3T3 transfectants, but not in association with mitotic spindle microtubules (stained with monoclonal antibody 20C6, green). **d-f**, Immunolocalization of wild-type survivin (red) in MSB-fixed NIH3T3 transfectants in association with mitotic microtubules in metaphase (**d**), midbodies in late telophase (**e**), and the MTOC in interphase (**f**, arrow). In **d-f**, the actin cytoskeleton was stained in green by FITC-phalloidin, and survivin was stained in red using a Texas-red-conjugated second antibody. **g**, Survivin inhibits apoptosis of NIH3T3 transfectants induced by the microtubule-stabilizing drug Taxol but is ineffective against the microtubule-depolymerizing drugs nocodazole and vincristine. Transfection of Bcl-2 is cytoprotective against all three drugs. Inset, inhibition of apoptosis in NIH3T3 transfectants by survivin, pcDNA3 or Bcl-2 at increasing concentrations of vincristine. For **a**, **g**, data are the mean $\pm$ s.e.m. of three independent experiments. For statistical analysis of *bcl-2* versus *survivin* transfectants, $P < 0.05$ was considered significant. * $P < 0.01$; *** $P < 0.0001$.

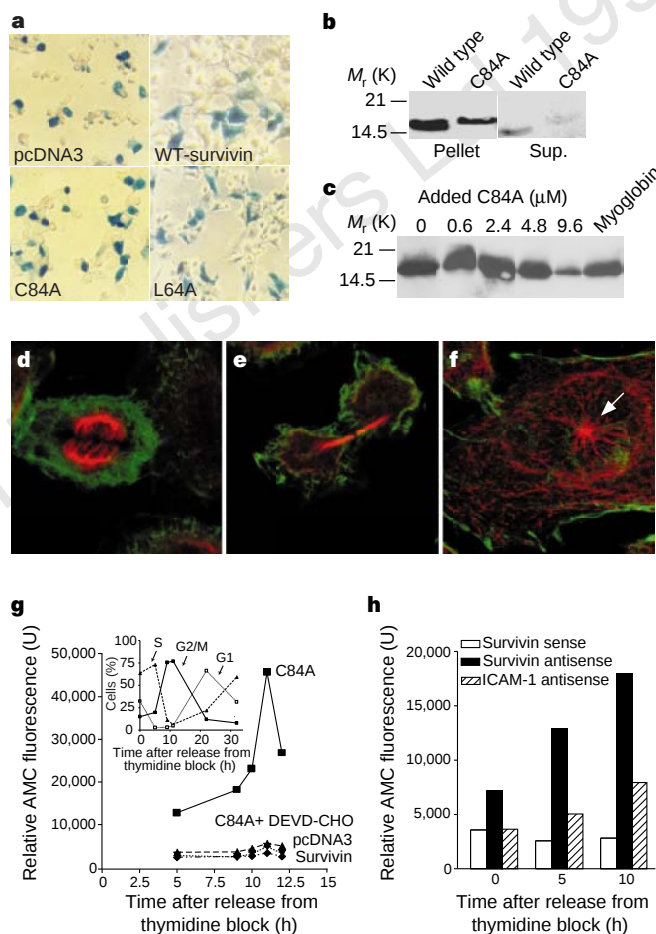


Figure 4 Disruption of survivin-microtubule interactions induces increased caspase-3 activity in G2/M phase. **a**, A C84A survivin BIR mutant fails to inhibit Taxol-induced apoptosis in β -galactosidase-expressing (blue) NIH3T3 transfectants, whereas wild-type (WT) survivin or the L64A survivin BIR mutant are cytoprotective, as judged by cell morphology. pcDNA3 is a control vector. **b**, the C84A BIR mutant binds to polymerized microtubules indistinguishably from wild-type survivin and is comparably depleted from supernatants (Sup.) of Taxol-treated samples. **c**, Biotinylated wild-type survivin is displaced from Taxol-stabilized microtubules by increasing concentrations of the C84A BIR mutant, but not by the control protein, myoglobin. **d-f**, Immunolocalization of the C84A BIR mutant (red) in MSB-fixed NIH3T3 transfectants in association with mitotic microtubules in metaphase (**d**), with midbodies in late telophase (**e**), and with the MTOC in interphase (**f**, arrow). In **d-f**, the actin cytoskeleton was stained in green with FITC-phalloidin, and survivin was stained in red using a Texas-red-conjugated second antibody. **g**, Displacement of endogenous survivin by forced expression of the C84A BIR mutant induces caspase-3-dependent DEVD-AMC cleavage in G2/M-synchronized HeLa cells. Cells with a subdiploid DNA content, as shown by propidium iodide staining, were 6.2% (pcDNA3), 7.7% (wild-type survivin) and 21.7% (C84A BIR mutant). DEVD-CHO is a caspase-3 inhibitor. Inset, cell-cycle analysis of DNA content of HeLa cells released from a thymidine-induced G1/S block. **h**, Antisense-mediated downregulation of endogenous survivin in HeLa cells²⁴ induces caspase-3-dependent DEVD cleavage in G2/M, whereas sense survivin RNA or ICAM-1 antisense RNA²⁴ are ineffective.

Leu 64 → Ala mutations were ineffective (Fig. 4a and data not shown). In co-sedimentation experiments, the C84A survivin BIR mutant bound to polymerized microtubules indistinguishably from wild-type survivin, and was similarly depleted from the supernatants of Taxol-treated samples (Fig. 4b). Increasing concentrations of the C84A BIR mutant effectively displaced biotinylated wild-type survivin from polymerized microtubules in a dose-dependent manner, whereas control myoglobin protein was ineffective (Fig. 4c). In immunofluorescence studies of NIH3T3 transfectants, the C84A BIR mutant bound to mitotic spindle microtubules in metaphase (Fig. 4d) and late telophase in midbodies (Fig. 4e), and was localized to the MTOC in interphase cells (Fig. 4f, arrow), indistinguishably from wild-type survivin (Fig. 3d–f). Under these conditions, displacement of endogenous survivin from microtubules by forced expression of the C84A BIR mutant resulted in a progressive increase in caspase-3 activity in synchronized HeLa cells (Fig. 4g). Maximal hydrolysis of the amino-acid sequence DEVD, a caspase-3 target, occurred 10–12 hours after release of cells from a G1/S thymidine-induced block, when ~80% of HeLa cells were in G2/M as shown by DNA-content analysis (Fig. 4g, inset); DEVD hydrolysis was suppressed by the caspase-3 inhibitor DEVD-CHO (Fig. 4g).

Lastly, we asked whether targeting endogenous survivin by antisense expression could also affect viability in G2/M. Forced expression of a survivin antisense construct downregulated expression of endogenous survivin in HeLa cells and resulted in increased caspase-3-dependent DEVD hydrolysis in G2/M-synchronized cells (Fig. 4h). In contrast, expression of a control antisense construct targeted towards intercellular adhesion molecule-1 (ICAM-1) was ineffective (Fig. 4h).

Our results identify survivin as a new interface between cell-cycle progression and control of apoptosis¹. Through its G2/M-specific expression, which is regulated transcriptionally as for typical mitotic genes⁹, and its topography, which is reminiscent of that of microtubule-associated proteins⁸, survivin may be required to counteract a constitutive pathway that induces apoptosis during mitosis. Like other IAP proteins^{11,13}, survivin inhibits the terminal effectors caspase-3 and caspase-7 (ref. 14). Caspase-3 has been implicated in proteolysis of structural proteins of the mitotic spindle during apoptosis¹⁵, and in co-localization with spindle fibres (our unpublished observations), indicating that survivin may preserve the integrity of the mitotic apparatus *in situ*. From the pattern of expression of survivin during development¹⁶, it seems possible that this apoptotic pathway may act as a general regulator of mitosis and may cooperate with other components of the G2/M checkpoint¹⁷ to preserve genetic fidelity during cell division^{18,19}. However, the overexpression of survivin in cancer⁷ may obliterate this apoptotic checkpoint and allow aberrant progression of transformed cells through mitosis. □

Methods

Cell-cycle synchronization. HeLa cells (ATCC) were treated with 400 μ M mimosine (to arrest cells in G1), 2 mM thymidine (to arrest cells in S) or 0.4 μ g ml⁻¹ nocodazole (to arrest cells in G2/M) (Sigma), for 16 h at 37 °C (ref. 20), or were arrested at the G1/S boundary by sequential culture with 2 mM thymidine and 400 μ M mimosine.

Proteins and antibodies. Wild-type or mutated survivin complementary DNAs were expressed as recombinant proteins fused to glutathione-S-transferase (GST)²¹. A truncated survivin coil-less mutant, consisting of residues M1–G99, was generated by the polymerase chain reaction (PCR). Mutagenesis of the survivin BIR was carried out with the Altered Site II system (Promega). The XIAP cDNA¹¹ was obtained by reverse-transcription and PCR-mediated amplification of Jurkat T-cell RNA. The anti-survivin monoclonal antibody 8E2 and the affinity-purified antibody towards the survivin sequence A3–I19 were as described^{7,21}.

Survivin promoter activity. Survivin–luciferase promoter constructs pLuc-230 (–39 to –268), pLuc-cyc1B (–36 to –268) and pLuc-cyc1.2 (+1 to –268)

(numbering from the initiating methionine) were generated by PCR and confirmed by DNA sequencing. HeLa cells were transiently transfected by LipofectAMINE, synchronized and analysed for luciferase activity in a Lumat LB 9510 luminometer. Values were normalized for β -galactosidase activity at A₄₀₅.

Immunofluorescence and confocal microscopy. HeLa cells or NIH3T3 fibroblasts (ATCC) transfected with wild-type survivin, survivin coil-less or the C84A BIR mutant in pcDNA3 (Invitrogen) were treated with Taxol (10 μ M) or colchicine (0.2 μ g ml⁻¹), methanol/acetone-fixed, and stained with anti-survivin monoclonal antibody 8E2 or anti- β -tubulin monoclonal antibody 20C6. Hoechst 33342 (2 μ g ml⁻¹) was used for chromosome staining. Non-binding mouse or rabbit IgG was used as a control. In some experiments methanol/acetone fixation was replaced by the MSB procedure²², which preserves cell structure. Coverslips were analysed on a Zeiss Axiophot microscope or by confocal laser scanning microscopy (CLSM Bio-Rad 1024).

Survivin–microtubule interaction. HeLa cell extracts (200 μ l, 4 mg ml⁻¹) or brain-purified tubulin (1 mg ml⁻¹, ICN Biochemicals) in MES buffer plus 1 mM GTP were incubated with 20 μ M Taxol (Sigma) or 40 μ M nocodazole (Sigma) for 30 min at 37 °C. After addition of wild-type or mutant survivin (0.6–3 μ M) for 30 min at 22 °C, samples were centrifuged through a 1-ml 10% sucrose cushion and samples from pellets or supernatants were analysed by Coomassie blue staining or immunoblotting with the antibody to the survivin sequence A3–I19 (ref. 7). We studied the reversibility of survivin–tubulin interactions during microtubule dynamics⁸ in successive rounds of temperature-dependent microtubule assembly/disassembly²³. The affinity of survivin for polymerized microtubules (K_d) was determined as the concentration of tubulin required to bind 50% of total added survivin¹⁰. For competition experiments, 0.3 μ M biotinylated wild-type survivin (sulphosuccinimidyl-6-(biotinamido) hexanoate; Pierce) was incubated with Taxol-stabilized polymerized microtubules in the presence of 0.6–9.6 μ M C84A survivin BIR mutant or myoglobin before co-sedimentation and detection by horseradish-peroxidase-conjugated streptavidin.

Apoptosis assay. NIH3T3 fibroblasts were transfected with 1 μ g *lacZ* and 4 μ g of various cDNA constructs by LipofectAMINE, and were treated with 1–5 μ M taxol, 0.8 μ g ml⁻¹ nocodazole or 1–100 μ g ml⁻¹ vincristine (Sigma). After a 16–24 h culture, β -galactosidase-expressing cells were scored morphologically as dead or alive⁶. Alternatively, HeLa cell transfectants were assayed at various time intervals after release from a G1/S boundary for caspase-3-dependent hydrolysis of the fluorogenic substrate Ac-DEVD-AMC (N-acetyl-Asp-Glu-Val-Asp-aldehyde; Pharmingen), in the presence or absence of the caspase-3 inhibitor Ac-DEVD-CHO. Fluorescence emissions were quantified on a spectrofluorometer with excitation wavelength of 360 nm and emission of 460 nm.

Received 5 October; accepted 30 October 1998.

1. Evan, G. I., Brown, L., Whyte, M. & Harrington, E. Apoptosis and the cell cycle. *Curr. Opin. Cell Biol.* **7**, 825–834 (1995).
2. Jacobson, M. D., Weil, M. & Raff, M. C. Programmed cell death in animal development. *Cell* **88**, 347–354 (1997).
3. Huang, D. C. S., O'Reilly, L. A., Strasser, A. & Cory, S. The anti-apoptosis function of bcl-2 can be genetically separated from its inhibitory effect on cell cycle entry. *EMBO J.* **16**, 4628–4638 (1997).
4. Minn, A. J., Boise, L. H. & Thompson, C. B. Expression of Bcl-X_L and loss of p53 can cooperate to overcome a cell cycle checkpoint induced by mitotic spindle damage. *Genes Dev.* **10**, 2621–2631 (1996).
5. Liston, P. *et al.* Suppression of apoptosis in mammalian cells by NAIP and a related family of IAP genes. *Nature* **379**, 349–353 (1996).
6. Uren, A. G., Pakusch, M., Hawkins, C. J., Puls, K. L. & Vaux, D. L. Cloning and expression of apoptosis inhibitory protein homologs that function to inhibit apoptosis and/or bind tumor necrosis factor receptor-associated factors. *Proc. Natl Acad. Sci. USA* **93**, 4974–4978 (1996).
7. Ambrosini, G., Adida, C. & Altieri, D. C. A novel anti-apoptosis gene, *survivin*, expressed in cancer and lymphoma. *Nature Med.* **3**, 917–921 (1997).
8. Hirokawa, N. Microtubule organization and dynamics dependent on microtubule-associated proteins. *Curr. Opin. Cell Biol.* **6**, 74–81 (1994).
9. Zwicker, J. *et al.* Cell cycle regulation of the cyclin A, cdc25C and cdc2 genes is based on a common mechanism of transcriptional repression. *EMBO J.* **14**, 4514–4522 (1995).
10. Butner, K. A. & Kirschner, M. W. Tau protein binds to microtubules through a flexible array of distributed weak sites. *J. Cell Biol.* **115**, 717–730 (1991).
11. Deveraux, Q. L., Takahashi, R., Salvesen, G. S. & Reed, J. X-linked IAP is a direct inhibitor of cell-death proteases. *Nature* **388**, 300–304 (1997).
12. Reed, J. C. Double identity for proteins of the Bcl-2 family. *Nature* **387**, 773–776 (1997).
13. Roy, N., Deveraux, Q. L., Takahashi, R., Salvesen, G. S. & Reed, J. C. The c-IAP-1 and c-IAP-2 proteins are direct inhibitors of specific caspases. *EMBO J.* **16**, 6914–6925 (1997).
14. Tamm, I. *et al.* IAP-family protein survivin inhibits caspase activity and apoptosis induced by Fas (CD95), Bax, and anticancer drugs. *Cancer Res.* (in the press).
15. Andrade, F. *et al.* Granzyme B directly and efficiently cleaves several downstream caspase substrates: implications for CTL-induced apoptosis. *Immunity* **8**, 451–460 (1998).

16. Adida, C. *et al.* Developmentally regulated expression of the novel cancer anti-apoptosis gene *survivin* in human and mouse differentiation. *Am. J. Pathol.* **152**, 43–49 (1998).
17. Murray, A. Cell cycle checkpoints. *Curr. Opin. Cell Biol.* **6**, 872–876 (1994).
18. Cross, S. M. *et al.* A p53-dependent mouse spindle checkpoint. *Science* **267**, 1353–1356 (1995).
19. Cahill, D. P. *et al.* Mutations of mitotic checkpoint genes in human cancers. *Nature* **392**, 300–303 (1998).
20. Hwang, A., Maity, A., McKenna, W. G. & Muschel, R. J. Cell cycle-dependent regulation of the cyclin B1 promoter. *J. Biol. Chem.* **270**, 28419–28424 (1995).
21. Lu, C.-D., Altieri, D. C. & Tanigawa, N. Expression of a novel antiapoptosis gene, *survivin*, correlated with tumor cell apoptosis and p53 accumulation in gastric carcinomas. *Cancer Res.* **58**, 1808–1812 (1998).
22. Osborn, M. & Weber, K. The display of microtubules in transformed cells. *Cell* **12**, 561–571 (1977).
23. Borisy, G. G., Marcum, J. M., Olmsted, J. B., Murphy, D. B. & Johnson, K. A. Purification of tubulin and associated high molecular weight proteins from porcine brain and characterization of microtubule assembly *in vitro*. *Ann. NY Acad. Sci.* **253**, 107–132 (1975).
24. Ambrosini, G., Adida, C., Sirugo, G. & Altieri, D. C. Induction of apoptosis and inhibition of cell proliferation by *survivin* gene targeting. *J. Biol. Chem.* **273**, 11177–11182 (1998).

Acknowledgements. We thank A. Villa and D. Adinolfi for assistance, and M. Osborn and K. Weber for monoclonal antibody 20C6. This work was supported by the NIH/NCI and the American Heart Association (D.C.A.), the Leukemia Research Foundation (G.A.), and Telethon and Associazione Italiana Ricerca sul Cancro (P.C.M.).

Correspondence and requests for materials should be addressed to D.C.A. (e-mail: dario.altieri@yale.edu).

Calcium promotes cell survival through CaM-K kinase activation of the protein-kinase-B pathway

Shigetoshi Yano, Hiroshi Tokumitsu & Thomas R. Soderling

Vollum Institute, Oregon Health Sciences University, Portland, Oregon 97201, USA

The protection against apoptosis provided by growth factors in several cell lines is due to stimulation of the phosphatidylinositol-3-OH kinase (PI(3)K) pathway, which results in activation of protein kinase B^{1,2} (PKB; also known as c-Akt and Rac) and phosphorylation and sequestration to protein 14-3-3 of the pro-apoptotic Bcl-2-family member BAD^{3–7}. A modest increase in intracellular Ca²⁺ concentration also promotes survival of some cultured neurons^{8,9} through a pathway that requires calmodulin but is independent of PI(3)K and the MAP kinases^{10,11}. Here we report that Ca²⁺/calmodulin-dependent protein kinase kinase (CaM-KK) activates PKB directly, resulting in phosphorylation of BAD on serine residue 136 and the interaction of BAD with protein 14-3-3. Serum withdrawal induced a three- to fourfold increase in cell death of NG108 neuroblastoma cells, and this apoptosis was largely blocked by increasing the intracellular Ca²⁺ concentration with NMDA (N-methyl-D-aspartate) or KCl or by transfection with constitutively active CaM-KK. The effect of NMDA on cell survival was blocked by transfection with dominant-negative forms of CaM-KK or PKB. These results identify a Ca²⁺-triggered signalling cascade in which CaM-KK activates PKB, which in turn phosphorylates BAD and protects cells from apoptosis.

To identify a Ca²⁺-mediated, PI(3)K-independent cell-survival mechanism, we tested whether CaM-KK might phosphorylate the 'activation-loop' residue T308 in PKB^{12–15}—CaM-KK phosphorylates similar activation-loop sites in calmodulin-dependent protein kinases I and IV (refs 16–18). CaM-KK alone exhibited strong *in vitro* autophosphorylation¹⁹ on multiple sites whereas a fusion protein containing glutathione-S-transferase and PKB exhibited little autophosphorylation; incubation of CaM-KK and GST-PKB together produced phosphorylation by CaM-KK of the PKB (Fig. 1a). Analysis of ³²P-labelled amino acids showed that the phosphorylated residue of PKB was a threonine (data not shown), presumably T308 because the T308A mutant was not phosphorylated by CaM-KK (Fig. 1a).

Phosphorylation of T308 should result in activation of PKB^{12–15}. α-CaM-KK or PKB alone each exhibited low kinase activity towards

histone H2B, but *in vitro* incubation of these proteins together resulted in a synergistic increase in kinase activity towards histone H2B (Fig. 1b). The T308A PKB mutant, which has normal basal PKB activity, showed little enhancement of activity in response to α-CaM-KK, indicating that CaM-KK was probably activating PKB. We confirmed this using a two-step assay: haemagglutinin (HA)-tagged PKB was expressed in COS-7 cells, immunoprecipitated and incubated for 20 min without or with recombinant α-CaM-KK, Ca²⁺/calmodulin and Mg²⁺/ATP. After washing the immune complex to remove CaM-KK, we studied PKB activity by assaying histone phosphorylation in the presence of EGTA, which inhibits any residual CaM-KK. CaM-KK catalysed strong activation of wild-type PKB in a Ca²⁺/calmodulin-dependent manner (Fig. 1c). Phosphorylation and activation of the T308A PKB mutant did not occur (Fig. 1a–c) and activation of wild-type PKB was reversed by treatment with protein phosphatase 2A (Fig. 1d), showing that activation of PKB by CaM-KK was due to phosphorylation, consistent with Fig. 1a. Expressed HA-PKB showed significant basal phosphorylation of its other regulatory site, S473 (ref. 15), determined using a phospho-specific antibody. Phosphorylation of S473 in HA-PKB was not increased by co-transfected CaM-KK activated by ionomycin but was markedly increased by treatment with platelet-derived growth factor (PDGF) (data not shown). The results of Fig. 1 show directly that recombinant CaM-KK can activate recombinant PKB *in vitro* through phosphorylation of T308.

Can CaM-KK activate PKB in cells? We co-transfected COS-7 cells with HA-PKB and either wild-type or constitutively active CaM-KK. In the absence of CaM-KK there was low PKB activity in the immune complex assay described above, but upon co-transfection with constitutively active CaM-KK (Fig. 2a) or wild-type CaM-KK plus ionomycin treatment (Fig. 2b) there was marked activation of wild-type PKB but not of the T308A mutant (Fig. 2a). Activation of PKB by CaM-KK was not sensitive to the PI(3)K inhibitor wortmannin (Fig. 2b, c). The rate of activation (Fig. 2c) was dependent on the amount of CaM-KK plasmid used (data not shown), and activation was reversed by subsequent *in vitro* treatment with protein phosphatase 2A (Fig. 2d).

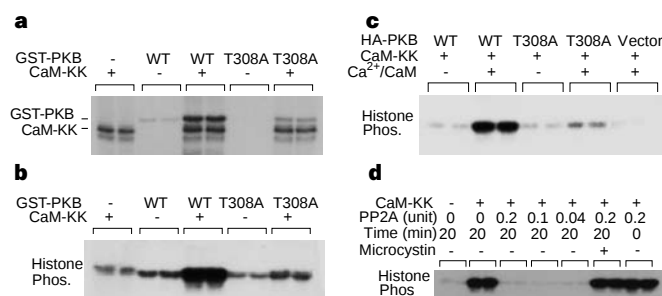


Figure 1 Phosphorylation and activation of recombinant PKB by CaM-KK *in vitro*. Each experiment in Figs 1 and 2 was replicated at least three times. **a**, GST-PKB (wild-type, WT, or T308A; 0.6 μg plasmid) was ³²P-labelled by CaM-KK (72 nM). Reaction products were analysed by SDS-PAGE/autoradiography. Duplicate samples are shown. **b**, GST-PKB (WT or T308A; 0.6 μg) was phosphorylated by CaM-KK (72 nM) in the presence of the PKB substrate histone H2B (0.2 mg ml⁻¹) and ³²P-ATP before analysis by SDS-PAGE/autoradiography. **c**, Immunoprecipitated HA-PKB (WT or T308A) was phosphorylated by CaM-KK (0.36 μM) and washed to remove CaM-KK, and the immune complex was assayed for PKB activity using histone H2B as a substrate in the presence of 2 mM EGTA, a Ca²⁺ chelator. **d**, HA-PKB (WT) was activated by CaM-KK and washed as in **c**. The indicated units of protein phosphatase 2A (PP2A) catalytic subunit were incubated for 0 or 20 min in the absence or presence of 1 μM of the PP2A inhibitor microcystin. The immune complex was washed again, and PKB activity was assayed, using histone H2B as a substrate, in the presence of 1 μM microcystin. CaM, calmodulin.

COS-7 cells do not contain detectable endogenous CaM-KK, so ionomycin treatment in the absence of transfected CaM-KK did not result in PKB activation (Fig. 2c), as shown previously for 3T3 cells²⁰. Therefore, in further experiments we used the neuroblastoma NG108 cell line, which contains endogenous CaM-KK²¹ and Ca²⁺-permeable NMDA-type ionotropic glutamate receptors. Stimulation of NG108 cells with NMDA or by depolarization with KCl resulted in a three- to fourfold activation of transfected or endogenous PKB that peaked at 1 hour after stimulation and remained elevated for up to 6 hours (Fig. 3a). The activation of HA-PKB by NMDA was blocked by dominant-negative PKB (Fig. 3c) or addition of EGTA (20 mM) or the NMDA antagonist MK-801 (10 μ M) to the medium (data not shown), and was apparently achieved by endogenous CaM-KK, as it was blocked by co-transfection with a dominant-negative CaM-KK (Fig. 3b). Activation of HA-PKB (sevenfold) by insulin-like growth factor (IGF), which was inhibited by wortmannin, was not blocked by dominant-negative CaM-KK (data not shown). HA-PKB activation by NMDA at 30 min was roughly twofold in either the presence (Fig. 3a) or the absence (Fig. 3b) of wortmannin, indicating that this is a PI(3)K-independent pathway.

PKB activation by CaM-KK should result in phosphorylation of BAD^{3,4} on residue S136 and its sequestration through interaction with protein 14-3-3 (refs 22, 23). We tested the ³²P-phosphorylation of GST-BAD *in vitro*; CaM-KK was not a catalyst, but PKB activated by CaM-KK was effective (data not shown). When NG108 cells were transfected with GST-BAD and then either transfected with constitutively active CaM-KK (Fig. 3d) or treated with NMDA (Fig. 3e), the phosphorylation of S136 in BAD was enhanced; the effect of NMDA was blocked by dominant-negative CaM-KK (Fig. 3e).

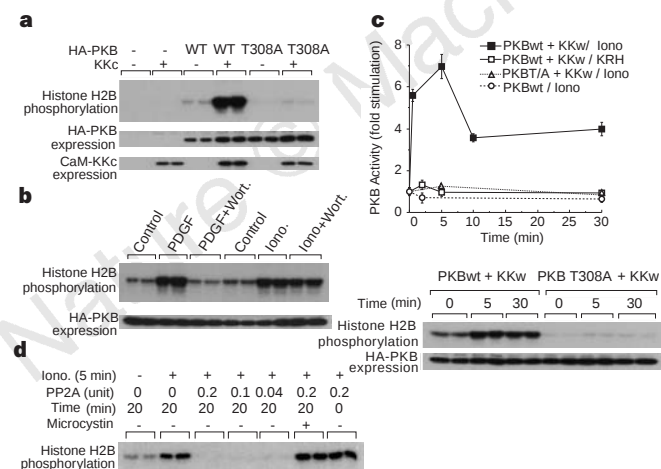


Figure 2 Activation of PKB by CaM-KK in transfected COS-7 cells. **a**, Cells were transfected with HA-PKB (WT or T308A; 0.6 μ g plasmid) without or with constitutively active CaM-KK (KKc) (0.6 μ g). HA-PKB was immunoprecipitated and assayed for activity with the substrate histone H2B (top). Western blots confirmed the expression levels of HA-PKB (centre) and CaM-KK (bottom). **b**, Cells transfected with HA-PKB (WT; 0.6 μ g) and CaM-KK (WT; 0.6 μ g) were treated for 10 min with PDGF or ionomycin with or without wortmannin. Immunoprecipitated PKB was assayed for activity (top) and level of expression (bottom). **c**, Cells were transfected with the indicated constructs (PKBT/A is the T308A PKB mutant; KKw is wild-type CaM-KK), and treated with KRH buffer containing 200 nM wortmannin without or with 1 μ M ionomycin. Top, immunoprecipitated HA-PKB was assayed and normalized to time zero; centre, autoradiograph of a typical experiment including ionomycin; bottom, levels of HA-PKB expression. **d**, Cells transfected with HA-PKB and CaM-KK were treated for 5 min with 1 μ M ionomycin. HA-PKB was immunoprecipitated and incubated with the indicated units of PP2A as in Fig. 1d; PKB activity towards histone H2B was measured.

Lastly, when endogenous BAD from NMDA-treated cells was immunoprecipitated, there was a time-dependent increase in co-precipitation of protein 14-3-3 (Fig. 3f).

We next studied apoptosis in NG108 cells upon serum withdrawal, using co-expressed green fluorescent protein (GFP) to identify cells transfected with the proteins mentioned below. Serum withdrawal for 24 hours induced a three- to fourfold increase in apoptosis; this enhanced apoptosis was blocked by 60–70% by treatment with Ca²⁺-mobilizing agonists that activate PKB, such as NMDA (Fig. 4a) or 25 mM KCl (data not shown). The protective effect of NMDA was blocked by transfection with dominant-negative PKB (Fig. 4a), supporting the idea that PKB is involved in this pathway. These results indicate that apoptosis of these cells in

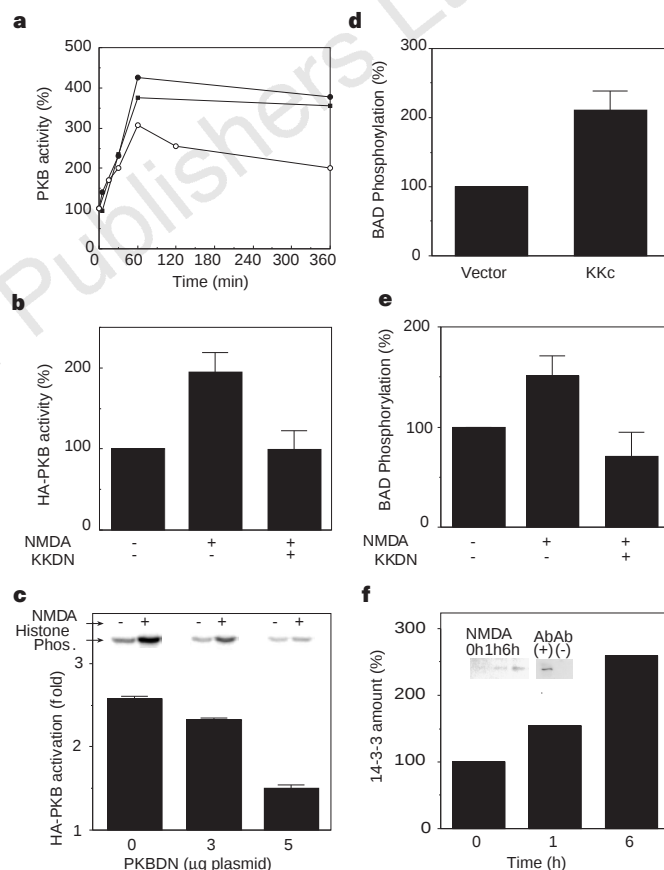


Figure 3 Phosphorylation of PKB and BAD as a result of CaM-KK activity in NG108 cells. **a**, Endogenous PKB (filled symbols) or transfected HA-PKB (1 μ g plasmid; open circles) was immunoprecipitated and assayed from NG108 cells treated with 100 μ M NMDA (circles) or 25 mM KCl (squares) in Mg²⁺-free KRH buffer containing 200 nM wortmannin. **b, c**, NG108 cells, transfected with HA-PKB (1 μ g plasmid) and either vector alone or vector expressing dominant-negative CaM-KK (KKDN; **b**; 5 μ g plasmid) or dominant-negative PKB (PKBDN; **c**), were treated with or without NMDA for 30 min; immunoprecipitated HA-PKB activity towards histone H2B was assayed. The inset in **c** shows one experiment $\pm$ NMDA. **d, e**, Cells transfected with GST-BAD (1 μ g plasmid) and either empty vector, constitutively active CaM-KK (KKc; **d**; 1 μ g plasmid) or KKDN (**e**; 5 μ g plasmid) were either non-stimulated (**d**) or treated for 1 h with or without NMDA (**e**). GST-BAD was precipitated and subjected to western analysis using an antibody specific for phosphorylated residue S136. **f**, Endogenous BAD was immunoprecipitated after treatment of cells with NMDA for 0, 1 or 6 h. The immune complex was subjected to western analysis of co-precipitated protein 14-3-3 (measured relative to time zero). Bars show averages from two experiments. Inset, the left three lanes show the results obtained from one experiment; the right two lanes show specific co-precipitation of 14-3-3 only, when anti-BAD antibody (Ab) was used for immunoprecipitation.

response to serum deprivation can be largely reversed by a Ca^{2+} -dependent, PKB-mediated pathway that may involve CaM-KK. If true, co-transfection with constitutively active CaM-KK should mimic the effect of NMDA treatment whereas dominant-negative CaM-KK should block the protective effect of NMDA. These were exactly the results observed (Fig. 4b). However, protection against apoptosis by treatment with IGF, which activates a PI(3)K-dependent pathway, was not reversed by dominant-negative CaM-KK (data not shown). In Fig. 4c, apoptosis of GFP-positive cells expressing dominant-negative CaM-KK (white dots with arrows) was about the same with (bottom panel) or without (middle panel) NMDA treatment, whereas in non-transfected cells (those not expressing GFP or dominant-negative CaM-KK) NMDA markedly protected against apoptosis (compare white dots without arrows in middle and bottom panels).

Programmed neuronal cell death is important in vertebrate neural development, during which roughly half of all neurons die. Survival usually results from the presence of neurotrophic factors provided by the target tissue or from persistent neuronal activity.

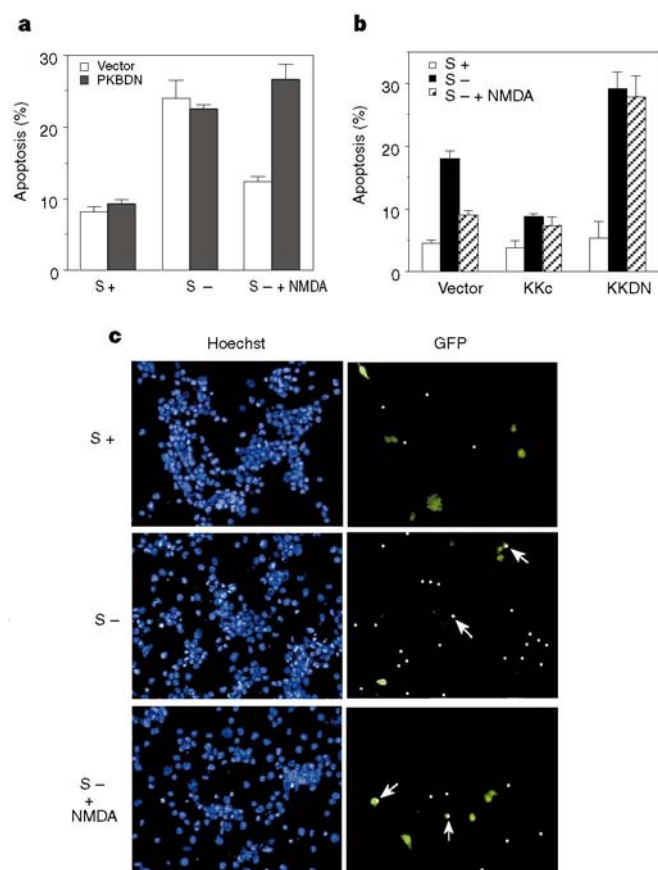


Figure 4 Regulation of apoptosis in NG108 cells. **a, b**, Cells transfected with GFP and either vector alone or vector expressing constitutively active CaM-KK (KKc; 1 μg), dominant-negative PKB (PKBDN) (5 μg) or dominant-negative CaM-KK (KKDN) (5 μg) were maintained in serum for 24 h (S+) or were serum-starved for 24 h in the absence (S-) or presence (S- NMDA+) of NMDA. GFP-positive cells were analysed for apoptosis by Hoechst staining: small, bright, condensed nuclei indicate apoptotic cells (see **c**). Apoptosis was quantified for each transfection/treatment condition in three plates using 150–200 GFP-positive cells. **c**, Selected fields of cells transfected with GFP (0.6 μg) and KKDN (5 μg) and treated as indicated (from the experiment giving the data shown in the right-hand three bars of **b**). Apoptotic cells were identified by Hoechst staining (left) and marked on computer images with a white dot, which was then overlaid with GFP immunofluorescence (right). Apoptotic GFP-positive cells are denoted by arrows.

For example, blockage of electrical activity in some types of developing neuron reduces their chance of survival during the period of programmed death²⁴. We propose a molecular model in which neuronal activity, by increasing the intracellular Ca^{2+} concentration to a modest degree, can protect cells against apoptosis. The increased intracellular Ca^{2+} concentration would stimulate CaM-KK, which in turn activates PKB. Activated PKB phosphorylates BAD^{3,4}, resulting in its sequestration by protein 14-3-3 (refs 22, 23). These results also indicate a broader role for CaM-KK, as it can phosphorylate and activate a cellular target protein (PKB) that is not also activated by Ca^{2+} /calmodulin, as is the case for calmodulin-dependent protein kinases I and IV (refs 16–18).

Methods

General. Protein kinase B (α -isoform) complementary DNA^{25,26} was cloned from rat brain messenger RNA using reverse transcription with polymerase chain reaction, and the T308A mutant was made by standard *in vitro* mutagenesis (5Prime $\rightarrow$ 3Prime, Inc.) and sequenced to confirm the mutation. The HA-tagged dominant-negative PKB double mutant (K179A (refs 1, 7) plus T308A) was catalytically inactive when expressed in COS-7 cells and assayed in the immune complex. The expression plasmids for wild-type CaM-KK and constitutively active CaM-KK have been described²¹. For *in vitro* experiments, CaM-KK was expressed in Sf9 cells and purified on CaM-Sepharose. Dominant-negative CaM-KK contained a K157A mutation. For transfections of COS-7 and NG108 cells, we used LipofectAMINE²⁷. HA-PKB was immunoprecipitated with monoclonal antibody 12CA5 (Boehringer) and endogenous PKB in NG108 cells was immunoprecipitated with an antibody against the pleckstrin-homology domain (Upstate Biotechnology). For experiments in which HA-PKB was activated by CaM-KK in COS-7 cells, transfected cells were pre-incubated for 1 h in Krebs–Ringer HEPES (KRH) containing 128 mM NaCl, 5 mM KCl, 1 mM Na_2HPO_4 , 1.2 mM MgSO_4 , 2.7 mM CaCl_2 , 10 mM glucose and 20 mM HEPES, pH 7.4, with or without 200 nM wortmannin for the last 10 min, and then treated with 20 ng ml^{-1} PDGF or 1 μM ionomycin in KRH for the indicated time, frozen and lysed. For western blotting of HA-PKB and CaM-KK, 10 μg cell lysate was subjected to SDS-PAGE and probed with 1,000-fold-diluted 12CA5 antibody (Boehringer) or anti-CaM-KK antibody (Transduction Laboratories), respectively. All treatments with NMDA also contained 1 μM glycine.

PKB phosphorylation and activation. *In vitro* activation and phosphorylation of glutathione-Sepharose-purified GST-PKB or immunoprecipitated HA-PKB by recombinant CaM-KK was done for 20–30 min with 50 mM HEPES, pH 7.5, 10 mM magnesium acetate, 1 mM dithiothreitol (DTT), 1 mM CaCl_2 , 3 μM calmodulin, 400 μM ^{32}P -ATP (1,000 c.p.m. pmol^{-1}) and baculovirus-expressed CaM-KK. Reactions were terminated by 20-fold dilution²⁷. PKB activity was then assayed for 30 min in 50 mM HEPES, pH 7.5, 10 mM magnesium acetate, 1 mM DTT, 1 μM protein kinase A inhibitor peptide, 400 μM ^{32}P -ATP (1,000 c.p.m. pmol^{-1}) and 0.2 mg ml^{-1} histone H2B; reaction products were separated by SDS-PAGE. Treatment of activated PKB with protein phosphatase 2A (ref. 28) was for 20 min. CaM-KK did not phosphorylate GST.

BAD phosphorylation. NG108 cells were transfected with GST-BAD (pEBG-MBad; New England Biolabs), serum-starved for 6 h and then stimulated with NMDA for the indicated times. Precipitated GST-BAD was subjected to western blotting with phospho-BAD (S136) antibody (Upstate Biotechnology), and values were normalized to the total amount of expressed BAD using general BAD antibody (Transduction Laboratories). Data expressed are from four to five independent experiments. For co-precipitation of 14-3-3 protein with BAD, non-transfected NG108 cells stimulated with 100 μM NMDA for the indicated times were lysed; endogenous BAD was then immunoprecipitated with 2 μg general BAD antibody (goat IgG; Santa Cruz). The immune complex was subjected to western blotting with anti-14-3-3 antibody (rabbit IgG; Upstate Biotechnology).

NG108 cell experiments and apoptosis. NG108-15 cells were cultured on 35-mm dishes²¹. For experiments with dominant-negative CaM-KK, cells were transfected with 1 μg HA-PKB plasmid and 5 μg plasmid containing dominant-negative CaM-KK or vector plasmid. To study apoptosis, we transfected NG108 cells on 35-mm dishes with 0.6 μg of GFP plasmid (pEGFP-C1;

Clontech) and either 1 μ g pcDNA3 (Invitrogen) vector expressing constitutively active CaM-KK or 5 μ g vector expressing dominant-negative CaM-KK or dominant-negative PKB (in all cases, equal amounts of vector alone were used for controls). The next day, media were changed to DMEM plus 10% serum, DMEM alone or DMEM containing 100 μ M NMDA plus 1 μ M glycine. After 24 h, cells were fixed with 4% paraformaldehyde for 15 min, stained for 15 min with 1 μ g ml⁻¹ Hoechst 33258 (CalBiochem) to visualize the nucleus, and stored in 50% glycerol in PBS at 4 °C. Apoptotic cells were visualized at 350 nm as small, bright, condensed nuclei, and GFP-positive cells were observed at 475 nm with a fluorescence microscope. Five fields from each dish were chosen randomly and captured in video image. The percentage of GFP-positive cells that were apoptotic was determined from three independent dishes per treatment condition. Typically 10–15% of cells were transfected, and 150–200 GFP-positive cells were scored for each treatment.

Received 2 September; accepted 29 October 1998.

1. Franke, T. F. *et al.* The protein kinase encoded by Akt proto-oncogene is a target of the PDGF-activated phosphatidylinositol 3-kinase. *Cell* **81**, 727–736 (1995).
2. Burgering, B. M. T. & Coffey, P. J. Protein kinase B (c-Akt) in phosphatidylinositol-3-OH kinase signal transduction. *Nature* **376**, 599–602 (1995).
3. Datta, S. R. *et al.* Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery. *Cell* **91**, 231–241 (1997).
4. Pesco, L., Gonzalez-Garcia, M., Page, C., Herrera, R. & Nunez, G. Interleukin-3-induced phosphorylation of BAD through the protein kinase Akt. *Science* **278**, 687–689 (1997).
5. Hemmings, B. A. Akt signaling: linking membrane events to life and death decisions. *Science* **275**, 628–630 (1997).
6. Marte, B. M. & Downward, J. PKB/Akt: connecting phosphoinositide 3-kinase to cell survival and beyond. *Trends Biochem. Sci.* **22**, 355–358 (1997).
7. Dudek, H. *et al.* Regulation of neuronal survival by the serine-threonine protein kinase Akt. *Science* **275**, 661–664 (1997).
8. Gallo, V., Kingsbury, A., Balazs, R. & Jorgensen, O. S. The role of depolarization in the survival and differentiation of cerebellar granule cells in culture. *J. Neurosci.* **7**, 2203–2213 (1987).
9. Koike, T., Martin, D. P. & Johnson, E. M. Role of calcium channels in the ability of membrane-depolarization to prevent neuronal death induced by trophic factor deprivation: evidence that levels of internal Ca²⁺ determine nerve growth factor dependence of sympathetic ganglion cells. *Proc. Natl Acad. Sci. USA* **86**, 6421–6425 (1989).
10. Hack, N., Hidaka, H., Wakefield, M. J. & Balazs, R. Promotion of granule cell survival by high K⁺ or excitatory amino acid treatment and Ca²⁺/calmodulin-dependent protein kinase activity. *Neuroscience* **57**, 90–120 (1993).
11. Soler, R. M. *et al.* Calmodulin is involved in membrane depolarization-mediated survival of motoneurons by phosphatidylinositol-3 kinase- and MAPK-independent pathways. *J. Neurosci.* **18**, 1230–1239 (1998).
12. Alessi, D. R. *et al.* Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase B α . *Curr. Biol.* **7**, 261–269 (1997).
13. Stokoe, D. *et al.* Dual role of phosphatidylinositol-3,4,5-trisphosphate in the activation of protein kinase B. *Science* **277**, 567–570 (1997).
14. Stephens, L. *et al.* Protein kinase B kinases that mediate phosphatidylinositol 3,4,5-trisphosphate-dependent activation of protein kinase B. *Science* **279**, 710 (1998).
15. Alessi, D. R. *et al.* Mechanism of activation of protein kinase B by insulin and IGF-1. *EMBO J.* **15**, 6541–6551 (1996).
16. Hariababu, B. *et al.* Human calcium-calmodulin dependent protein kinase I: cDNA cloning, domain structure and activation by phosphorylation at threonine-177 by calcium-calmodulin dependent protein kinase I kinase. *EMBO J.* **14**, 3679–3686 (1995).
17. Seibert, M. A. *et al.* Phosphorylation and activation of Ca(2+)-calmodulin-dependent protein kinase IV by Ca(2+)-calmodulin-dependent protein kinase Ia kinase. Phosphorylation of threonine 196 is essential for activation. *J. Biol. Chem.* **270**, 17616–17621 (1995).
18. Tokumitsu, H., Enslin, H. & Soderling, T. R. Characterization of a CaM-kinase cascade: molecular cloning and characterization of CaM-kinase. *J. Biol. Chem.* **270**, 19320–19324 (1995).
19. Wayman, G. A., Tokumitsu, H. & Soderling, T. R. Inhibitory cross-talk by cAMP-kinase on the calmodulin-dependent protein kinase cascade. *J. Biol. Chem.* **272**, 16073–16076 (1997).
20. Conus, N. M., Hemmings, B. A. & Pearson, R. B. Differential regulation by calcium reveals distinct signaling requirements for the activation of akt and p70. *J. Biol. Chem.* **273**, 4776–4782 (1998).
21. Enslin, H., Tokumitsu, H., Stork, P. J. S., Davis, R. J. & Soderling, T. R. Regulation of MAP kinases by a calcium/calmodulin-dependent protein kinase cascade. *Proc. Natl Acad. Sci. USA* **93**, 10803–10808 (1996).
22. Zha, J., Harada, H., Yang, E., Jockel, J. & Korsmeyer, S. J. Serine phosphorylation of death agonist BAD in response to survival factor results in binding to 14-3-3 not Bcl-X_L. *Cell* **87**, 619–628 (1996).
23. Hsu, S. Y., Kaipia, A., Zhu, L. & Hsueh, A. J. Interference of BAD (Bcl-2-associated death promoter)-induced apoptosis in mammalian cells by 14-3-3 isoforms and P11. *Mol. Endocrinol.* **11**, 1858–1867 (1997).
24. Furbre, S., Oppenheim, R. W. & Prevette, D. Naturally-occurring neuron death in the ciliary ganglion of the chick embryo following removal of preganglionic input: evidence for the role of afferents in ganglion cell survival. *J. Neurosci.* **7**, 1816–1832 (1987).
25. Coffey, P. J. & Woodgett, J. R. Molecular cloning and characterization of a novel putative protein-serine kinase related to the cAMP-dependent and protein kinase C families. *Eur. J. Biochem.* **201**, 475–481 (1991).
26. Jones, P. F., Jakubowicz, T., Pitossi, F. J., Maurer, F. & Hemmings, B. A. Molecular cloning and identification of a serine/threonine protein kinase of the second-messenger subfamily. *Proc. Natl Acad. Sci. USA* **88**, 4171–4175 (1991).
27. Tokumitsu, H. & Soderling, T. R. Requirements for calcium and calmodulin in the calmodulin-kinase activation cascade. *J. Biol. Chem.* **271**, 5617–5622 (1996).
28. Park, J. K. & Soderling, T. R. Activation of Ca²⁺/calmodulin-dependent protein kinase IV by CaM-kinase kinase in Jurkat T lymphocytes. *J. Biol. Chem.* **270**, 30464–30469 (1995).

Acknowledgements. We thank B. Chang for help in constructing dominant-negative PKB and P. Stork for discussions during the course of this work. This study was supported by the NIH.

Correspondence and requests for materials should be addressed to T.R.S. (e-mail: soderlit@ohsu.edu).

Decoupling of nucleotide- and microtubule-binding sites in a kinesin mutant

Hebok Song & Sharyn A. Endow

Department of Microbiology, Duke University Medical Center, Durham, North Carolina 27710, USA

Molecular motors require ATP to move along microtubules or actin filaments. To understand how molecular motors function, it is crucial to know how binding of the motor to its filamentous track stimulates the hydrolysis of ATP by the motor, enabling it to move along the filament. A mechanism for the enhanced ATP hydrolysis has not been elucidated, but it is generally accepted that conformational changes in the motor proteins^{1–3} occur when they bind to microtubules or actin filaments, facilitating the release of ADP. Here we report that a mutation in the motor domain of the microtubule motor proteins Kar3 and Ncd uncouples nucleotide- and microtubule-binding by the proteins, preventing activation of the motor ATPase by microtubules. Unlike the wild-type motors, the mutants bind tightly to both ADP and microtubules, indicating that interactions between the nucleotide- and microtubule-binding sites are blocked. The region of the motor that includes the mutated amino acid could transmit or undergo a conformational change required to convert the motor ATPase into a microtubule-stimulated state.

Kar3, a minus-end-directed kinesin motor of *Saccharomyces cerevisiae*, is essential in karyogamy (nuclear fusion) following mating and in meiosis, and has a non-essential role in mitosis^{4–6}. Kar3 is thought to counteract two plus-end-directed kinesin motors, Cin8 and Kip1, in the mitotic spindle^{6,7}. The mutants recovered in a screen for suppressors of a *cin8 kip1* double mutant

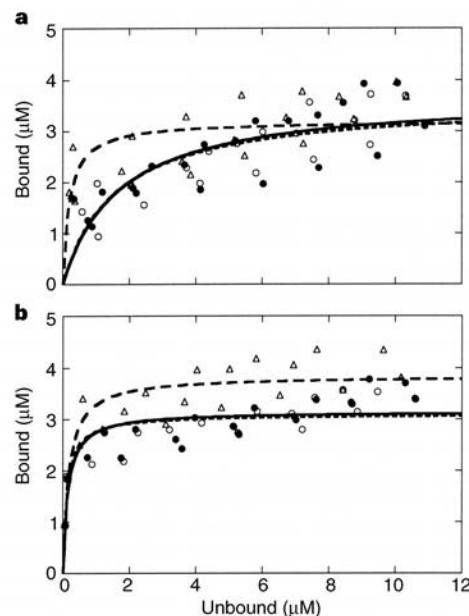


Figure 1 Microtubule binding by wild-type Kar3 and Kar3-898. Microtubule-binding assays were done without added nucleotide or in the presence of 1 mM Mg-ADP or 1 mM Mg-ATP. Data from three separate experiments for each nucleotide condition were combined and a hyperbolic curve was fitted to the data points. **a**, Kar3 and **b**, Kar3-898 protein. Open triangles and curves comprising long dashes, no added nucleotide; open circles and solid curves, Mg-ADP; filled circles and curves with short dashes, Mg-ATP.

all mapped to the Kar3 motor domain⁸. One mutant, *kar3-898* (in which residue N650 was converted to K), is a dominant missense suppressor of the *cin8-ts kip1-Δ* double mutant. Not only has *kar3-898* lost its wild-type activity *in vivo*, but it is also dominant in combination with wild-type Kar3 (ref. 8).

We cloned the DNA encoding the Kar3-898 motor domain and purified⁹ the mutant protein; we then studied microtubule binding by Kar3-898 and the wild-type Kar3 protein⁹. We estimated the K_d values (dissociation constants) for Kar3 from microtubules to be $1.7 \pm 0.7 \mu\text{M}$ ($n = 3$) and $1.6 \pm 0.6 \mu\text{M}$ ($n = 3$) in the presence of ADP and ATP, respectively, whereas those of Kar3-898 were $0.12 \pm 0.04 \mu\text{M}$ ($n = 3$) and $0.14 \pm 0.04 \mu\text{M}$ ($n = 3$) (Fig. 1). In the absence of nucleotide, the K_d values were $0.19 \pm 0.08 \mu\text{M}$ ($n = 3$) for Kar3 and $0.17 \pm 0.04 \mu\text{M}$ ($n = 3$) for Kar3-898. These data show that binding of Kar3-898 to microtubules is insensitive to nucleotide. In contrast to wild-type Kar3, Kar3-898 binds tightly to microtubules in the presence or absence of ADP or ATP, yielding K_d values that were the same as for the wild-type protein in the absence of nucleotide. The measurable K_d values show that the binding is not irreversible.

We also determined whether the Kar3-898 protein could bind to and hydrolyse ATP. When $[\alpha\text{-}^{32}\text{P}]\text{ATP}$ was incubated with the wild-type or mutant protein and the reaction mix was separated on a gel-filtration column, the proteins co-eluted with a peak of radioactivity (Fig. 2a, b). When $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ was incubated with the proteins, the proteins separated from the peak of radioactivity (Fig. 2c, d). The

co-elution of radioactivity with the proteins after incubation with $[\alpha\text{-}^{32}\text{P}]\text{ATP}$ shows that ATP binds to both proteins, whereas the separation of radioactivity from the proteins incubated with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ shows that ATP is hydrolysed by both proteins and that P_i is released. These results show that the Kar3-898 protein retains ATP-binding and -hydrolytic activities.

We then tested the ability of microtubules to stimulate the ATPase activity of Kar3-898. In a steady-state ATPase assay using a coupled enzyme system¹⁰, the basal ATPase activity of the mutant protein was similar to that of the wild-type protein, that is, only slightly above background in these assays. Strikingly, unlike that of the wild-type protein, the ATPase activity of the mutant motor was not stimulated by up to $2 \mu\text{M}$ microtubules (Fig. 3a). Because the steady-state assay is relatively insensitive to low ATPase activity, we carried out further experiments using mant-ATP, a fluorescent ATP analogue; we monitored the release of bound mant-ADP from the protein upon addition of excess unlabelled ATP. The rate constant for mant-ADP release from Kar3 was $0.006 \pm 0.0004 \text{ s}^{-1}$ ($n = 6$) in the absence of microtubules (Fig. 3b). When microtubules were added with the unlabelled ATP, the dissociation rate of mant-ADP from the wild-type protein increased with increasing microtubule concentration ($0.022 \pm 0.005 \text{ s}^{-1}$, $n = 4$, for $0.25 \mu\text{M}$ microtubules; $0.037 \pm 0.008 \text{ s}^{-1}$, $n = 4$, for $0.5 \mu\text{M}$ microtubules) (Fig. 3b). The rate of mant-ADP release from Kar3-898 in the absence of microtubules ($0.004 \pm 0.0003 \text{ s}^{-1}$, $n = 4$) (Fig. 3c) was only slightly

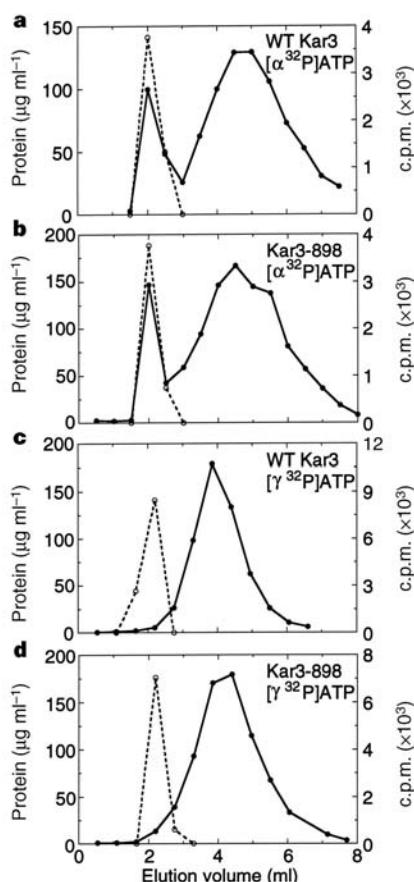


Figure 2 ATP binding and hydrolysis by wild-type Kar3 and Kar3-898. Kar3-898 was tested in parallel with Kar3 to determine whether the mutant protein retains the ability to bind and hydrolyse ATP. The proteins were incubated with $[\alpha\text{-}^{32}\text{P}]\text{ATP}$ or $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ and free radioactivity was separated from protein by gel-filtration chromatography. **a**, **b**, Kar3 (**a**) and Kar3-898 (**b**) incubated with $[\alpha\text{-}^{32}\text{P}]\text{ATP}$. **c**, **d**, Kar3 (**c**) and Kar3-898 (**d**) incubated with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$. Protein, dashed line; radioactivity, solid line. WT, wild type.

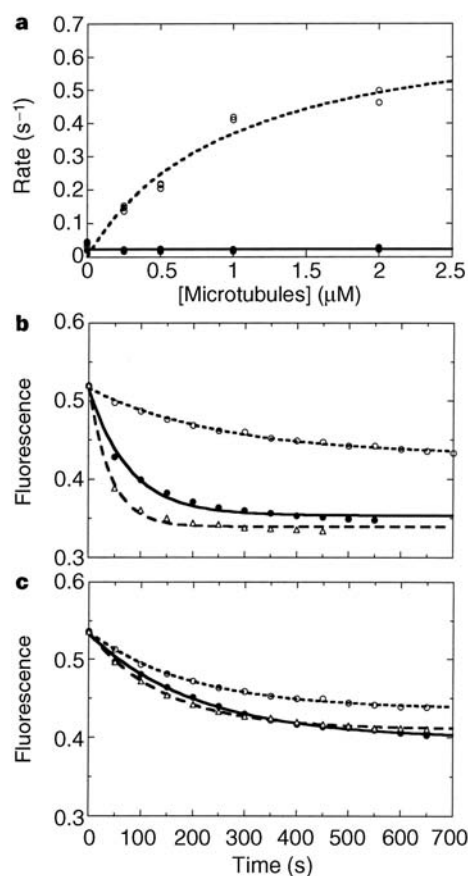


Figure 3 ATP hydrolysis and mant-ADP release by wild-type Kar3 and Kar3-898. **a**, Steady-state ATPase activity of the Kar3 and Kar3-898 proteins. Open circles, Kar3; filled circles, Kar3-898. **b**, Mant-ADP release from Kar3 after adding excess unlabelled ATP in the absence (open circles) or presence of microtubules ($0.25 \mu\text{M}$, filled circles, or $0.5 \mu\text{M}$, triangles). **c**, Mant-ADP release from Kar3-898 in the absence (open circles) or presence of microtubules ($0.5 \mu\text{M}$, filled circles, or $2.0 \mu\text{M}$, triangles). Fluorescence, relative units.

different from that of wild-type Kar3. However, there was little or no difference in the dissociation rate of mant-ADP from Kar3-898 in the presence of 0.5 μM ($0.004 \pm 0.0003 \text{ s}^{-1}$, $n = 6$) or 2.0 μM ($0.005 \pm 0.0003 \text{ s}^{-1}$, $n = 3$) microtubules compared with the absence of microtubules (Fig. 3c). The mutation in the Kar3 motor domain thus prevents stimulation of the motor to release ADP upon binding to microtubules. The Kar3-898 protein binds tightly to ADP in the presence of microtubules, whereas ADP release from wild-type Kar3 is accelerated by the binding of the motor to microtubules. The defect in Kar3-898 is not due to an inability to bind microtubules or hydrolyse ATP, but to a failure to convert the basal ATPase activity of the motor to the microtubule-stimulated state. The tight binding by Kar3-898 to both ADP and microtubules indicates a defect in interactions between the nucleotide- and microtubule-binding sites. Our data do not exclude the possibility of an effect of the mutation on ATP hydrolysis, but this does not alter the conclusion that interactions between the two sites are blocked in the mutant.

Residue N650 of Kar3 is invariant among the kinesin proteins¹¹. We therefore mutated the corresponding amino acid of Ncd, N600 to K, producing Ncd-NK. We purified⁹ and analysed Ncd-NK in parallel with the wild-type Ncd protein (MC6)¹². The K_d values for dissociation of wild-type Ncd from microtubules in the presence of ADP and ATP were estimated to be $5.3 \pm 0.8 \mu\text{M}$ ($n = 3$) and $5.8 \pm 1.1 \mu\text{M}$ ($n = 3$), respectively, whereas those for Ncd-NK were $3.0 \pm 0.6 \mu\text{M}$ ($n = 2$) and $1.8 \pm 0.4 \mu\text{M}$ ($n = 2$) under our assay conditions. The increased binding affinity of Ncd-NK for microtubules in the presence of ADP or ATP, compared with the binding

affinity of the wild-type protein, was not as great as that of Kar3-898, but nevertheless we did observe significantly lower K_d values, corresponding to tighter binding, for Ncd-NK than for Ncd. Nucleotide-binding experiments with [α -³²P]ATP and [γ -³²P]ATP resulted in the same radioactivity elution profiles for Ncd and Ncd-NK as for Kar3 and Kar3-898 (Fig. 2), showing that Ncd-NK can bind and hydrolyse ATP. In steady-state assays, the ATPase activity of wild-type Ncd was stimulated >100-fold by addition of 4 μM microtubules, but there was no stimulation of Ncd-NK ATPase activity at the same concentration of microtubules (Fig. 4a). The rate of mant-ADP release from Ncd was $0.002 \pm 0.0004 \text{ s}^{-1}$ ($n = 6$) in the absence of microtubules and $0.03 \pm 0.003 \text{ s}^{-1}$ ($n = 6$) in the presence of 0.2 μM microtubules (Fig. 4b). The rate of mant-ADP release from Ncd-NK was $0.002 \pm 0.0001 \text{ s}^{-1}$ ($n = 6$) in the absence of microtubules and nearly the same ($0.003 \pm 0.0002 \text{ s}^{-1}$, $n = 4$) in the presence of 2.0 μM microtubules (Fig. 4c), a tenfold higher concentration than that assayed for wild-type. The tighter binding of Ncd-NK to microtubules compared with Ncd, and the absence of microtubule-stimulated ATPase activity and mant-ADP release, parallel the results obtained for Kar3-898.

As Kar3 is slower than Ncd in motility assays¹³, we constructed a fusion protein containing glutathione S-transferase (GST) and Ncd-NK to test the effect of the mutation on microtubule motility. Microtubule-gliding assays were performed¹⁴ using GST-Ncd-NK in parallel with wild-type GST-Ncd (GST-MC1)¹². Microtubules bound to coverslips coated with GST-Ncd-NK lysates, but no microtubule movement was observed. The GST-Ncd-NK lysates contained as much or more soluble motor protein, estimated by western blot analysis, as control GST-Ncd lysates that showed active microtubule gliding. The binding of microtubules to the coverslip by GST-Ncd-NK but the absence of movement is consistent with our biochemical results showing that the Ncd-NK protein can bind to microtubules but the ATPase activity of the mutant protein is not stimulated by microtubules.

The Kar3-898 and Ncd-NK biochemical defects are caused by a single missense mutation present in helix $\alpha 4$ of Kar3 (ref. 15) and in the adjacent loop, L11, of Ncd¹⁶. The mutation separates the basal from the microtubule-stimulated ATPase activity of the motor by blocking interactions between the nucleotide- and microtubule-binding sites, implicating L11 and $\alpha 4$ in coupling the two sites. L11 is thought to be involved in microtubule binding by kinesin on the basis of alanine-scanning mutagenesis¹⁷. L11/ $\alpha 4$ may be analogous to the G-protein 'switch II' region, which changes in conformation upon nucleotide exchange¹⁶. In the kinesin proteins, the structural elements that include the mutated residue could transmit or undergo a conformational change on binding of the wild-type motor to microtubules, a change that fails to occur in Kar3-898 and Ncd-NK. Uncoupling mutants have been recovered previously in *Dictyostelium* myosin II (ref. 18). The S465V (S474 in chicken myosin S1) mutation maps to a loop/helix analogous to L11/ $\alpha 4$ of kinesin¹⁹, and may correspond to residues N650 and N600 in Kar3 and Ncd. The S465V mutant has an increased basal ATPase activity that is almost equal to that of the wild-type actin-activated ATPase, but shows a lower velocity than the wild-type protein¹⁸. These mutant effects differ from those of Kar3-898 and Ncd-NK, indicating that the myosin and kinesin motor mechanisms could differ considerably despite the apparent structural homology of the motors. From the crystal structure of the Ncd motor domain, it was proposed that L11 and $\alpha 4$ are involved in communication between the nucleotide- and microtubule-binding sites¹⁶. Our studies of Kar3-898 and Ncd-NK substantiate this prediction. □

Methods

Plasmids. A BglII–BseRI polymerase chain reaction (PCR) fragment prepared from *kar3-898* (ref. 8) DNA was ligated to BglII + BseRI-digested plasmid *pMW172/kar3* (ref. 9), encoding residues 383–729 of Kar3, to obtain the

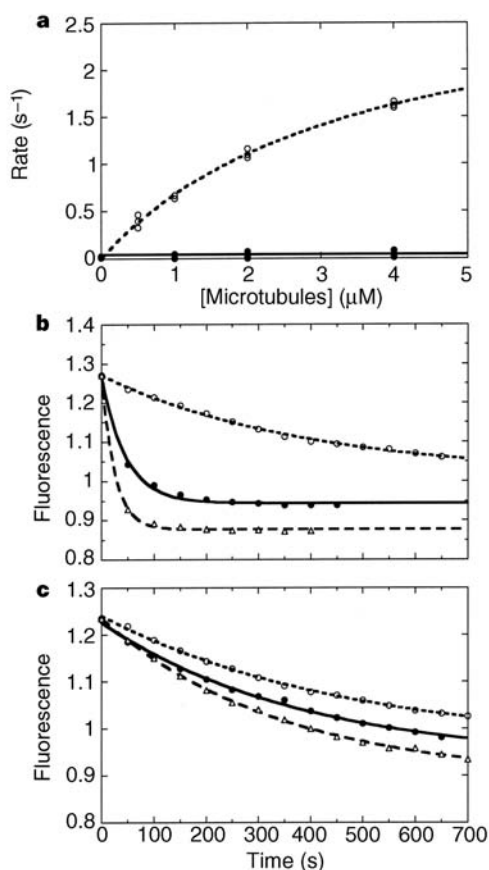


Figure 4 ATP hydrolysis and mant-ADP release by wild-type Ncd and Ncd-NK. **a**, Steady-state ATPase activity of wild-type Ncd and Ncd-NK. Open circles, Ncd; filled circles, Ncd-NK. **b**, Mant-ADP release from Ncd in the absence (open circles) or presence of microtubules (0.1 μM , filled circles, or 0.2 μM , triangles). **c**, Mant-ADP release from Ncd-NK in the absence (open circles) or presence of microtubules (0.5 μM , filled circles, or 2.0 μM , triangles).

plasmid *pMW172/kar3-898*. Analysis of the sequence between the *Bgl*III and *Bse*RI sites confirmed the mutation and the absence of PCR missense mutations. *pMW172/kar3-898* was transformed into *BL21(DE3)pLysS* host cells²⁰ for protein expression. Overlap extension PCR²¹ was used to construct *pMW172/ncd-NK*. The *Sac*I-digested PCR product was inserted into *Sac*I-digested *pMW172/ncd* (MC6)¹² which encodes residues 333–700 of the full-length protein. The N600K mutation was confirmed by DNA sequencing of the *Sac*I fragment. *GST-ncd-NK* was constructed by replacing the *Bst*EII + *Eco*RI fragment of *GST-MC1* (ref. 12) with the corresponding fragment of *pMW172/ncd-NK*. The mutated region of the fragment was sequenced to confirm the mutation.

Microtubule-binding assays. Assays were performed by incubating protein (1–14 μ M) with 3.5 μ M microtubules [tubulin heterodimer] for 5 min at room temperature²². 50 mM NaCl was added to the Kar3 and Kar3-898 reaction mixtures and 40 mM NaCl was added to the assays for Ncd and Ncd-NK. Protein bound to microtubules was separated from unbound protein by centrifugation at 100,000g for 15 min at 22°C. The supernatants and pellets were analysed by SDS–PAGE. Coomassie-blue-stained gels were scanned to product digital images and protein bands were quantified using NIH Image. The binding curves were obtained by plotting the amount of protein bound to microtubules against the level of unbound protein and fitting a hyperbolic curve to the data points using Kaleidagraph version 3.0.5.

Nucleotide-binding assays. Wild-type or mutant Kar3 or Ncd protein (10 μ M) was incubated on ice for 12–14 h with 90 μ M [α -³²P]ATP or [γ -³²P]ATP in a 200- μ l reaction mix containing 50 mM Tris-acetate, pH 7.5, 1 mM dithiothreitol and 1 mM Na₂EGTA. Unbound nucleotide was separated from protein on a Sephadex G-50 gel-filtration column. 10 μ l and 30 μ l of each fraction were withdrawn for radioactivity and protein determination, respectively. Radioactivity was counted on a Beckman LS 3801 liquid scintillation counter and the protein concentration of each fraction was measured using the Bio-Rad protein assay reagent.

ATPase assays. Steady-state ATPase activity was measured using a pyruvate kinase/lactate dehydrogenase-coupled NADH oxidation reaction in the presence or absence of microtubules¹⁰. For mant-ADP-release experiments, mant-ATP (4 μ M) was incubated with protein (2 μ M) for 2 h on ice before addition of unlabelled ATP plus or minus microtubules²³. Fluorescence intensity at 446 nm was recorded every 10 s for 800 s after addition of unlabelled ATP. Fluorescence was plotted at 50-s intervals and data points were fitted to a single exponential equation using Kaleidagraph version 3.0.5. As the initial drop in fluorescence intensity of the reaction mixtures could not be monitored while adding unlabelled ATP, we determined the zero time point values from parallel reactions in which buffer was added instead of unlabelled ATP. This was especially important for the reactions in which microtubules were added with the unlabelled ATP to wild-type Kar3 and Ncd.

Received 2 September; accepted 28 October 1998.

- Hirose, K., Lockhart, A., Cross, R. A. & Amos, L. A. Nucleotide-dependent angular change in kinesin motor domain bound to tubulin. *Nature* **376**, 277–279 (1995).
- Whittaker, M. et al. A 35-Å movement of smooth muscle myosin on ADP release. *Nature* **378**, 748–751 (1995).
- Dominguez, R., Frey, Y., Trybus, K. M. & Cohen, C. Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: visualization of the pre-power stroke state. *Cell* **94**, 559–571 (1998).
- Meluh, P. B. & Rose, M. D. KAR3, a kinesin-related gene required for yeast nuclear fusion. *Cell* **60**, 1029–1041 (1990).
- Bascom-Slack, C. A. & Dawson, D. S. The yeast motor protein, Kar3p, is essential for meiosis I. *J. Cell Biol.* **139**, 459–467 (1997).
- Saunders, W., Lengyel, V. & Hoyt, M. A. Mitotic spindle function in *Saccharomyces cerevisiae* requires a balance between different types of kinesin-related motors. *Mol. Biol. Cell* **8**, 1025–1033 (1997).
- Saunders, W. S. & Hoyt, M. A. Kinesin-related proteins required for structural integrity of the mitotic spindle. *Cell* **70**, 451–458 (1992).
- Hoyt, M. A., He, L., Totis, L. & Saunders, W. S. Loss of function of *Saccharomyces cerevisiae* kinesin-related *CIN8* and *KIP1* is suppressed by KAR3 motor domain mutations. *Genetics* **135**, 35–44 (1993).
- Song, H. & Endow, S. A. Rapid purification of microtubule motor domain proteins expressed in bacteria. *BioTechniques* **22**, 82–85 (1997).
- Huang, T.-G. & Hackney, D. D. *Drosophila* kinesin minimal motor domain expressed in *Escherichia coli*. *J. Biol. Chem.* **269**, 16493–16501 (1994).
- Bloom, G. S. & Endow, S. A. Motor proteins I: kinesins. *Protein Prof.* **2**, 1109–1171 (1995).
- Chandra, R., Salmon, E. D., Erickson, H. P., Lockhart, A. & Endow, S. A. Structural and functional domains of the *Drosophila* ncd microtubule motor protein. *J. Biol. Chem.* **268**, 9005–9013 (1993).
- Endow, S. A. et al. Yeast Kar3 is a minus-end microtubule motor protein that destabilizes microtubules preferentially at the minus ends. *EMBO J.* **13**, 2708–2713 (1994).
- Song, H., Golovkin, M., Reddy, A. S. N. & Endow, S. A. *In vitro* motility of AtKCBP, a calmodulin-binding kinesin protein of *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **94**, 322–327 (1997).
- Gulick, A. M., Song, H., Endow, S. A. & Rayment, I. X-ray crystal structure of the yeast Kar3 motor domain complexed with Mg-ADP to 2.3 Å resolution. *Biochemistry* **37**, 1769–1776 (1998).

- Sablin, E. P., Kull, F. J., Cooke, R., Vale, R. D. & Fletterick, R. J. Crystal structure of the motor domain of the kinesin-related motor ncd. *Nature* **380**, 555–559 (1996).
- Woehlke, G. et al. Microtubule interaction site of the kinesin motor. *Cell* **90**, 207–216 (1997).
- Ruppel, K. M. & Spudis, J. A. Structure-function studies of the myosin motor domain: importance of the 50-kDa cleft. *Mol. Biol. Cell* **7**, 1123–1136 (1996).
- Kull, F. J., Sablin, E. P., Lau, R., Fletterick, R. J. & Vale, R. D. Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. *Nature* **380**, 550–555 (1996).
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89 (1990).
- Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51–59 (1989).
- Song, H. & Endow, S. A. Binding sites on microtubules of kinesin motors of the same class: opposite polarity. *Biochemistry* **35**, 11203–11209 (1996).
- Moore, J. D., Song, H. & Endow, S. A. A point mutation in the microtubule-binding region of the Ncd motor protein reduces motor velocity. *EMBO J.* **15**, 3306–3314 (1996).

Acknowledgements. We thank A. Hoyt for *kar3-898*, C. Fierke and K. Hightower for help with fluorimetry and the use of a fluorometer, S. Rosenfeld for mant-ATP and A. Gulick for comments on the manuscript. This work was supported by the NIH.

Correspondence and requests for materials should be addressed to S.A.E. (e-mail: endow@galactose.mc.duke.edu).

Identification of the receptor component of the I κ B α -ubiquitin ligase

Avraham Yaron*†, Ada Hatzubai*†, Matti Davis*, Iris Lavon*, Sharon Amit*, Anthony M. Manning‡, Jens S. Andersen§, Matthias Mann§, Frank Mercurio‡ & Yinon Ben-Neriah*

* The Lautenberg Center for Immunology, The Hebrew University-Hadassah Medical School, Jerusalem 91120, Israel

‡ Signal Pharmaceuticals Inc., San Diego, California 92121, USA

§ Protana A/S, Staermosegaardsvej 16, DK-5230 Odense M, Denmark

† These authors contributed equally to this work

NF- κ B, a ubiquitous, inducible transcription factor involved in immune, inflammatory, stress and developmental processes, is retained in a latent form in the cytoplasm of non-stimulated cells by inhibitory molecules, I κ Bs^{1–3}. Its activation is a paradigm for a signal-transduction cascade that integrates an inducible kinase and the ubiquitin–proteasome system to eliminate inhibitory regulators. Here we isolate the pI κ B α -ubiquitin ligase (pI κ B α -E3) that attaches ubiquitin, a small protein which marks other proteins for degradation by the proteasome system, to the phosphorylated NF- κ B inhibitor pI κ B α . Taking advantage of its high affinity to pI κ B α , we isolate this ligase from HeLa cells by single-step immunoaffinity purification. Using nanoelectrospray mass spectrometry, we identify the specific component of the ligase that recognizes the pI κ B α degradation motif as an F-box/WD-domain protein belonging to a recently distinguished family of β -TrCP/Slimb proteins. This component, which we denote E3RS^{I κ B} (pI κ B α -E3 receptor subunit), binds specifically to pI κ B α and promotes its *in vitro* ubiquitination in the presence of two other ubiquitin-system enzymes, E1 and UBC5C, one of many known E2 enzymes. An F-box-deletion mutant of E3RS^{I κ B}, which tightly binds pI κ B α but does not support its ubiquitination, acts *in vivo* as a dominant-negative molecule, inhibiting the degradation of pI κ B α and consequently NF- κ B activation. E3RS^{I κ B} represents a family of receptor proteins that are core components of a class of ubiquitin ligases. When these receptor components recognize their specific ligand, which is a conserved, phosphorylation-based sequence motif, they target regulatory proteins containing this motif for proteasomal degradation.

Cytokine stimulation of different cell types results in the activation of the I κ B kinase (IKK) complex^{4–11}, leading to I κ B α phosphorylation at serine residues 32 and 36 and, consequently, ubiquitin-dependent degradation of pI κ B α (refs 12, 13). We previously observed that the pI κ B α from proteasome-inhibited cells is associated with cellular complexes that are distinguishable from non-phosphorylated I κ B α complexes by their large relative mole-

cular mass¹⁴. A probable explanation for the increased mass of the pIkB α complexes could be the recruitment of components of the ubiquitin machinery¹⁴. To explore this possibility, we purified pIkB α /NF- κ B complexes from proteasome-inhibited, tumour necrosis factor- α (TNF- α)-stimulated HeLa cells and studied their ubiquitination potential. After a short stimulation with TNF- α , IkB α /NF- κ B complexes were immunoaffinity-purified with anti-RelA (p65) antibodies and the cognate p65 peptide. The immunopurified fraction was supplemented with various components of the ubiquitin system and subjected to *in vitro* ubiquitination. We found that the addition of ubiquitin, purified E1 and a specific E2, UBC5C (purified independently as the IkB α -E2 (A.Y. *et al.*, unpublished results)), was sufficient to generate the full-capacity IkB α -ubiquitin conjugating activity (Fig. 1, lane 2), evident in the accumulation of high molecular mass species that reacted with IkB α -specific antisera. This activity was E1-dependent (compare lanes 1 and 2 in Fig. 1), and was provided only by immunopurified fractions from TNF α -stimulated HeLa cells (compare lanes 4–6 in Fig. 1). The addition of a phosphopeptide comprising the pIkB α degradation motif (pp10)¹⁴ (Fig. 1, lane 7), but not of a serine/glutamic-acid-substituted IkB α peptide (pS/E) (Fig. 1, lane 8), abolished the formation of polyubiquitin-IkB α conjugates, emphasizing the specificity of the conjugation process.

IkB α phosphorylation can be reproduced *in vitro* by a recombinant, constitutively active IKK- β protein (IKK2-EE)⁶. After immunoprecipitation of ³⁵S-labelled IkB α /NF- κ B complexes from a non-stimulated HeLa lysate previously incubated with recombinant ³⁵S-labelled IkB α , the complexes were phosphorylated by the recombinant IKK2-EE, eluted with the p65 cognate peptide and subjected to *in vitro* ubiquitination. After incubation with IKK2-EE, nearly all of the ³⁵S-labelled IkB α was phosphorylated. However, the addition of ubiquitin, E1 and UBC5C did not result in pIkB α

ubiquitination (Fig. 2a, lane 2). Therefore, IkB α phosphorylation by IKK was not sufficient to promote its ubiquitination in the presence of E1 and E2. Conceivably, pIkB α ubiquitination requires an additional component of the HeLa lysate that was not co-immunopurified from non-stimulated cells. To test this hypothesis, immuno-bound IkB α /NF- κ B complexes were incubated with a non-stimulated HeLa lysate, either directly or after IKK2-EE phosphorylation, washed extensively with high-salt buffer and eluted with the p65 peptide. Incubation of the phosphorylated IkB complexes (Fig. 2a, lane 3), but not of the non-phosphorylated ones (Fig. 2a, lane 1), with the HeLa lysate provided the pIkB-ligase component(s) necessary for pIkB α conjugation. No signal was obtained when E1 or E2 was omitted from the reaction, confirming that the signal at the top of the gel represents polyubiquitin-IkB α conjugates (Fig. 2a, lanes 6 and 7). The addition of pp10 (Fig. 2a, lane 4), but not of pS/E (Fig. 2a, lane 5), to the incubation mixture abolished the ubiquitin-ligase activity adsorbed to the pIkB complex, while preserving the integrity of the substrate. This was evident in the ability of the peptide-treated fractions to undergo ubiquitination in the presence of reticulocyte fraction II as an E3 source¹³ (data not shown). These results indicate that, as a result of IkB α phosphorylation, an essential component of the pIkB α -ubiquitin ligase is recruited from the HeLa lysate through a specific and direct interaction with the pIkB degradation motif¹⁴.

DS-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the IkB α /NF- κ B complex did not reveal any protein that may have been recruited after IKK phosphorylation (not shown). The complexity of the protein staining, however, could obscure the presence of any recruited protein migrating along with an immunopurified protein. To circumvent the possible masking of the putative pIkB-ubiquitin ligase, we used as a ligase source ³⁵S-biosynthetically labelled HeLa lysate and traced the pIkB α -associated proteins by

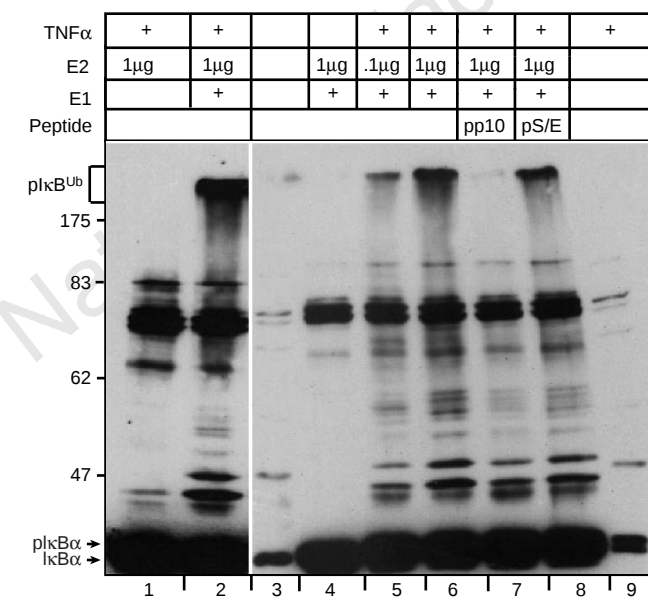


Figure 1 Cytokine stimulation promotes the association between pIkB α and a specific ubiquitin ligase. HeLa cells were pre-incubated with the proteasome inhibitor ALLN and stimulated with TNF- α for 10 min. The IkB α /NF- κ B complex was then immunoprecipitated from TNF α -stimulated or non-stimulated HeLa cell lysates using anti-RelA (see Methods) and subjected to *in vitro* ubiquitination upon addition of ubiquitin, ATP- γ S and the following components: lane 1, UBC5C; 2, UBC5C and E1; 3, a sample of non-stimulated HeLa lysate; 4–6, UBC5C and E1; 7, UBC5C, E1 and pIkB α peptide; 8, UBC5C, E1 and serine-substituted IkB α peptide; 9, a sample of TNF α -stimulated HeLa lysate. Cell stimulation is indicated in the TNF- α row. Monomeric IkB α , pIkB α -polyubiquitin conjugates and M_r size markers (K) are indicated on the left. The intense signals at 72–75K originate in crossreactive background proteins.

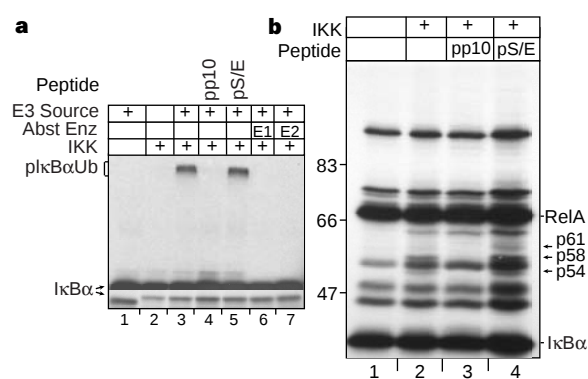


Figure 2 IkB α phosphorylation is necessary and sufficient to recruit specific ubiquitin-ligase activity. **a**, Ubiquitination assay. ³⁵S-labelled IkB α /NF- κ B complex immunopurified from non-activated cells was phosphorylated by IKK2-EE (IKK, where indicated), incubated with non-activated HeLa lysate as an E3 source (indicated), washed and subjected to *in vitro* ubiquitination in the presence of ATP- γ S, ubiquitin, E1, and UBC5C (except where an excluded component is indicated by Abst Enz). A pIkB α peptide (lane 4) or a serine-substituted IkB α peptide (lane 5) was added during HeLa incubation. A strip of a brief gel-exposure is shown at the bottom, to distinguish IkB α (lane 1) from pIkB α (lanes 2–7). **b**, Proteins adsorbed onto pIkB α /NF- κ B from ³⁵S-labelled HeLa cells. ³⁵S-radiolabelled HeLa lysate was incubated with antibody-immobilized IkB α /NF- κ B complex. The immune complexes were then washed, eluted and analysed by SDS-PAGE and autoradiography. Lane 1, non-phosphorylated IkB α /NF- κ B complex incubated with HeLa lysate; lanes 2–4, phosphorylated IkB α /NF- κ B complex incubated with HeLa lysate in the absence (lane 2) or presence of pIkB α peptide (lane 3) or serine-substituted IkB α peptide (lane 4). Indicated on the left are M_r size markers (K); on the right are the three pIkB α -associated bands that were displaced by the pIkB α peptide (arrows).

SDS-PAGE analysis and autoradiography (Fig. 2b). In parallel, the various fractions were tested for their ubiquitin-ligase activity (not shown). The band pattern of the active fraction (Fig. 2b, lane 2) was compared with that of the non-active one (Fig. 2b, lane 1). Four ^{35}S -labelled protein bands with relative molecular masses of 54,000, 58,000, 61,000 and 64,000 (M_r 54K, 58K, 61K and 64K) were distinguished in lane 2. To single out the ligase component that recognizes $\text{pI}\kappa\text{B}\alpha$ directly, pp10 or the control peptide pS/E was added to the radiolabelled HeLa lysate, which was then incubated with the immuno-bound $\text{I}\kappa\text{B}\alpha/\text{NF-}\kappa\text{B}$ complex. A comparison of the eluted fractions showed that of the four distinctive bands present only in lane 2, three bands were eliminated by the specific pp10 peptide (p54, p58 and p61), whereas only the 64K band persisted in the presence of pp10 (Fig. 2b, compare lanes 2 and 3). The control peptide did not affect the association of any of the distinctive proteins with $\text{pI}\kappa\text{B}\alpha$ (Fig. 2b, lane 4). Two of the $\text{pI}\kappa\text{B}\alpha$ interacting proteins, p58 and p54, were consistently present and always associated with the specific ubiquitin-ligase activity. The 54K and 58K bands were excised from a ligase-positive and a ligase-negative (a non-phosphorylated $\text{I}\kappa\text{B}\alpha$ complex incubated with HeLa lysate) lane, the proteins were digested *in situ*, and the tryptic peptides obtained were sequenced by nanoelectrospray mass spectrometry¹⁵. Mass spectra of the 54K gel band revealed a complex

peptide mixture (Fig. 3a), derived predominantly from NF- κB and $\text{I}\kappa\text{B}$ proteins, from which several peptides were selected for fragmentation. To identify the protein associated with the E3 activity, additional peptides, which were present in small amounts, were selected for sequencing by comparing the spectrum of the 54K bands from the active fraction (Fig. 3c) with that of a similar band from the non-active one (Fig. 3b). The peptide-sequence tag (1587.81) VVNV (1999.09) was derived from the fragmentation spectrum shown in Fig. 3c and identified unambiguously as AAVNVVDFDDKYIVSAS, a fragment of the human F-box/WD protein $\beta\text{-TrCP}$. The result was confirmed by further spectra, corresponding to the human $\beta\text{-TrCP}$ peptides VLEGHEELVR, LVVSGSSDNTIR, and IQDIETIESNWR¹⁶. Similar peptides were identified from the band corresponding to p58 (not shown). The ability to identify proteins present in small amounts, despite the presence of high levels of co-migrating proteins, illustrates how new mass-spectrometry techniques can be used for micro-characterization of gel-separated proteins. The high sensitivity afforded by nanoelectrospray quadrupole time-of-flight mass spectrometry, combined with the high data quality¹⁷, allowed extensive characterization of complex peptide mixtures and selection of distinct peptides for sequencing, and significantly improved the interpretation of the resulting fragmentation spectra.

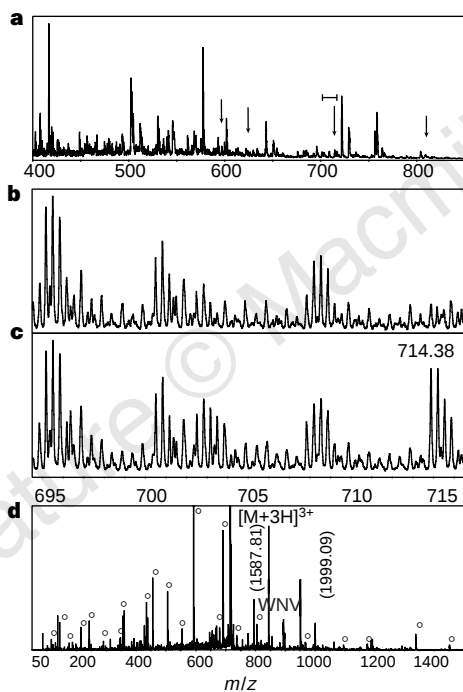


Figure 3 Identification of the E3-associated, $\text{pI}\kappa\text{B}\alpha$ -binding protein. **a**, Nanoelectrospray mass spectrum of the unseparated tryptic peptide mixture from the 54K gel band excised from a ligase-positive lane (marked in lane 2 in Fig. 2b). Peaks marked by arrows were fragmented and identified as peptides derived from $\beta\text{-TrCP}$. Fragmentation of the unmarked peaks identified the proteins NF- κB1 (p50), IKK- α , $\text{I}\kappa\text{B}\epsilon$, RelB, tubulin beta-1 chain, and thyroid receptor interacting protein 9. The bar indicates the region enlarged in **c**. **b**, **c**, Comparison of the nanoelectrospray spectra of the 54K band associated with (**c**) and without (**b**) ubiquitin-ligase activity. The peptide at mass-to-charge (m/z) 714.38, which was only present in the active fraction, was selected for sequencing. **d**, Fragmentation spectrum of the peptide identified in **c**. A sequence tag was assembled from a series of doubly charged fragment ions and searched in the NRDB for a matching pattern. Fragment masses calculated for the retrieved $\beta\text{-TrCP}$ sequence AAVNVVDFDDKYIVSASGDR were compared with the complete fragmentation spectrum to confirm the match. Peaks matching expected fragment ions are marked by circles. The height of an individual peak in each panel reflects its relative intensity.

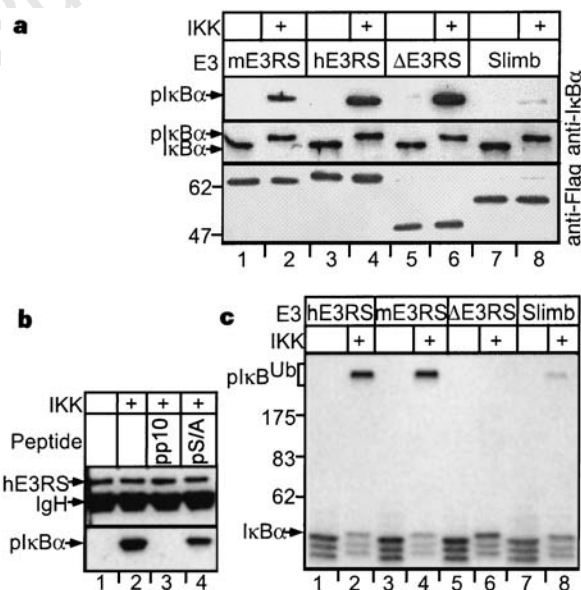


Figure 4 Binding and ubiquitination specificity of E3RS proteins. **a**, Selective binding to $\text{pI}\kappa\text{B}\alpha$. Proteins were immunoprecipitated through a Flag epitope from transfected 293T cells, incubated with immunopurified $\text{I}\kappa\text{B}\alpha/\text{NF-}\kappa\text{B}$ complex, which had been treated with or without IKK as indicated, and the bound material was analysed by western blotting with the indicated antibodies. The top panel shows specific $\text{pI}\kappa\text{B}\alpha$ binding; the middle panel shows 10% of the substrate flow-through; the bottom panel is a blot of the immunoprecipitated Flag-tagged E3RS binding proteins; M_r size markers (K) are indicated on the left. **b**, $\text{hE3RS}^{\text{K}\kappa\text{B}}$ - $\text{pI}\kappa\text{B}\alpha$ binding is abolished by a phosphopeptide representing the $\text{pI}\kappa\text{B}\alpha$ degradation motif (pp10)¹⁴, but not by a related non-phosphorylated peptide (pS/A). **c**, *In vitro* ubiquitination of $\text{pI}\kappa\text{B}\alpha$ by the E3RS proteins. Immunopurified Flag-tagged proteins were incubated with ^{35}S -labelled $\text{I}\kappa\text{B}\alpha/\text{NF-}\kappa\text{B}$ complexes, treated with IKK as indicated, and subjected to ubiquitination in the presence of ATP- γS , ubiquitin, E1 and UBC5C. The $\text{I}\kappa\text{B}\alpha$ substrate (the full-length product and two degradation products), $\text{pI}\kappa\text{B}\alpha$ -polyubiquitin conjugates and M_r size markers (K) are indicated on the left.

After identification of the β -TrCP protein, we isolated and characterized human, mouse and *Drosophila* β -TrCP-related complementary DNA clones (see Methods) and investigated their properties in cell-free and cell-based assays that report pIkB α ubiquitination. On the basis of its documented function (see below), we suggest renaming the human and mouse β -TrCP proteins hE3RS^{IKB} and mE3RS^{IKB}, respectively (relating to the E3 receptor subunit for I κ B, where substrate specificity is assigned by a superscript). To qualify as the substrate-recognition component of the I κ B ligase, an E3RS protein is expected to bind pIkB α specifically and assist in its ubiquitination. These two parameters were examined in a cell-free system. The NF- κ B/I κ B α complex was immunopurified from HeLa cells and the immune complex was either phosphorylated with IKK-2EE or mock-phosphorylated (see Methods). It was then incubated with the following immobilized Flag-tagged E3RS proteins immunoprecipitated from transfected 293T cells: mE3RS^{IKB}, hE3RS^{IKB}, Slimb, and the F-box-deletion mutant of hE3RS^{IKB}, Δ E3RS^{IKB}. The bound material was analysed by western blotting with anti-I κ B α and anti-Flag antibodies. All E3RS proteins exclusively bound IKK-phosphorylated but not a mock-phosphorylated I κ B α (Fig. 4a). However, the human and mouse E3RS^{IKB} bound pIkB α far more effectively than did the highly homologous *Drosophila* protein (compare lanes 2, 4, 6 and 8 in Fig. 4a). Δ E3RS^{IKB} bound pIkB α even better than the wild-type protein, indicating that the F-box region was dispensable for binding. Furthermore, E3RS^{IKB} binding was abolished by a peptide representing the pIkB-recognition motif (Fig. 4b, lane 3), but not by the control peptide (Fig. 4b, lane 4), specifying the site of pIkB α recognition as the conserved DS(PO₃)GLDS(PO₃) sequence.

The observed binding specificity suggested that E3RS^{IKB} would serve as a core protein of the I κ B α -ubiquitin ligase. To test its function, immunopurified wild-type and mutant Flag-tagged E3RS^{IKB} proteins were used as a source of E3 in pIkB α ubiquitination. In the presence of E1 and E2 (UBC5C), the wild-type E3RS^{IKB} proteins facilitated the ubiquitination of pIkB α , but not of the non-phosphorylated I κ B α (Fig. 4c, lanes 1–4). Δ E3RS^{IKB}, devoid of the F-box protein–protein interaction module, failed to promote ubiquitination (Fig. 4c, lane 6), in spite of its binding capacity (Fig. 4a, lane 6). Although Slimb facilitated some pIkB α ubiquitination, it was at least ten times less efficient than the human and mouse E3RS^{IKB} (based on similar Flag-tag expression levels), corresponding to its weaker binding activity. These experiments indicate a critical role for E3RS^{IKB} in controlling pIkB α ubiquitination.

The modular design of E3RS^{IKB} and our *in vitro* analysis suggested that Δ E3RS^{IKB}, which binds pIkB α but fails to promote its ubiquitination *in vitro*, would function as a dominant-negative molecule *in vivo*. Indeed, transient overexpression of Δ E3RS^{IKB} inhibited the degradation of endogenous I κ B α in stimulated Jurkat cells, resulting in accumulation of pIkB α (Fig. 5). As a consequence, activation of NF- κ B was inhibited (Fig. 5), and this inhibition was specific, as Δ E3RS^{IKB} did not affect the activation of an NF-AT reporter (data

not shown). Note that NF- κ B inhibition was also observed with wild-type Slimb, whereas the expression of wild-type hE3RS^{IKB} was not inhibitory (Fig. 5). Therefore, overexpression of wild-type Slimb has a dominant negative effect on NF- κ B activation, which is probably linked to its relatively poor pIkB α -ubiquitination activity (Fig. 4c).

Protein destruction, like protein synthesis, is a fundamental mechanism by which organisms control diverse cellular functions^{18–20}. We have elucidated a critical step in mobilizing a signalling cascade that couples signal-dependent phosphorylation to degradation of a regulatory protein; this step is pIkB α recognition, mediated by a subset of F-box/WD proteins that specialize as receptors for the degradation motif⁴. Recently, human β -TrCP was isolated as a protein interacting with Vpu, a small human immunodeficiency virus (HIV)-encoded protein that directs the CD4 protein of T cells to degradation¹⁶. β -TrCP seems to bind Vpu through a sequence similar to the DS(PO₃)G Ψ XS(PO₃) motif²¹, and was postulated to function as a ubiquitin ligase¹⁶. The latter case represents an intriguing example of how a pathogen adapts to its host. By assimilating a specific host-protein degradation motif, HIV recruits the host ubiquitination machinery to eliminate a cellular protein. However, there is a notable difference between the targeting of pIkB α and Vpu by E3RS^{IKB}: whereas pIkB α is recognized and targeted by the E3RS^{IKB} for ubiquitination, Vpu is solely the ligand, directing the E3 to an associated host protein, CD4. The only other known β -TrCP-family member, Slimb, was isolated by genetic screening, as an enhancer of the Wntless and Hedgehog signalling pathways in *Drosophila*²². Our *in vitro* experiments show that Slimb, despite being highly homologous to E3RS^{IKB}, is a rather poor E3 receptor for pIkB α and, in fact, behaves as a dominant-negative molecule in NF- κ B activation *in vivo*. In this respect, it is notable that Slimb mutations do not affect the Cactus/Dorsal system²², the *Drosophila* homologues of I κ B α /NF- κ B, indicating a remarkable substrate-specificity in the β -TrCP/Slimb family.

E3RS^{IKB} belongs to a large and growing family of F-box proteins, the two prototypes of which are the yeast proteins Cdc4 and Grr1 (refs 23–25), which share 30% similarity with E3RS^{IKB}. These proteins, which specialize in ubiquitin-dependent degradation of cell-cycle regulatory proteins, have, aside from an F-box, one of two distinct protein–protein interaction modules, WD or leucine-rich repeat (LRR) regions. These modules were shown recently to mediate the binding of the phosphorylated targets Sic1 and Cln2 to Cdc4 and Grr1, respectively^{23–25}. However, target binding to these proteins would be enhanced and might be effective only in the context of a complex termed SCF, where S stands for Skp1, C for cullin and F for the F-box protein^{23,26}. SCF complexes do not seem to have a catalytic function of their own, but rather direct a ubiquitin-conjugating enzyme (E2) towards a specific substrate, particularly after substrate phosphorylation^{23–25}. The modular design of these complexes suggests that a similar adapter system is associated with E3RS^{IKB}. However, so far we have failed to detect Skp1 or Cul1 in association either with the endogenous I κ B α -ligase complex, or with Flag-E3RS^{IKB} (data not shown). In principle, a ligase could comprise only two distinct subunits, receptor and E2, the latter functioning as the catalytic subunit. In this case, substrate-specific ubiquitination would simply be governed by association of a distinct receptor with a defined E2, a matter for future studies. □

Methods

Reagents, cDNA cloning and plasmids. Synthetic I κ B α peptide pp10 (DRHDS[PO₃]GLDS[PO₃]M), the serine-substituted peptides pS/E and pS/A and the p65 peptide (ELFPLIFPAEPAQASGP) were purchased from Alfa-Diagnostic Inc. Sepharose-immobilized and purified goat anti-p65 (sc-109 AC and sc-109, respectively) and rabbit anti-I κ B α (sc-371) antibodies were purchased from Santa Cruz. Monoclonal anti-Flag (M2) and the Flag-M2 peptide were purchased from Sigma. His₆-UBC5C (ref. 27) was expressed in BAL21 bacteria and purified on Ni²⁺-NTA Agarose (QIAGEN) according to the

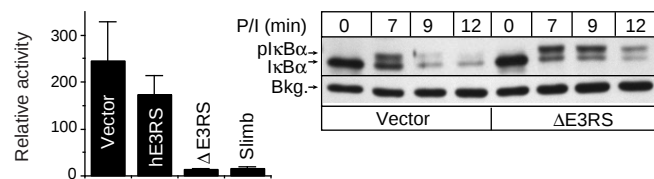


Figure 5 Inhibition of I κ B α degradation and NF- κ B activation by overexpression of Δ E3RS^{IKB}, a dominant-negative molecule. Right, western blot analysis of I κ B α of phorbol-ester and Ca²⁺ ionophore (P/I)-stimulated and non-stimulated Jurkat cells transfected with an empty Flag vector or Δ E3RS^{IKB}. The post-stimulation interval (min) is indicated. Left, κ B-dependent luciferase assay in P/I-stimulated Jurkat cells transfected with κ B-Luc reporter plasmid⁶ and the indicated expression vectors. NF- κ B activity is expressed as relative (fold) luciferase activity, the non-stimulated empty Flag vector being the reference (single-fold).

manufacturer's procedure. E1 was purified from HeLa cells¹⁴. Slimb was cloned from the total RNA of mature *Drosophila* flies by polymerase chain reaction (PCR) into CMV2-Flag at the *HindIII* site. hE3RS^{IKB} (a β -TrCP variant¹⁶ containing an in-frame insertion after amino acid 16 of the sequence CSMPSRLWLGCSSLADSMPSRLCLYNPGTGALTAFAQ) was cloned from a Jurkat cDNA library by PCR and inserted into CMV2-flag at the *NorI* site. Δ E3RS^{IKB} was derived from hE3RS^{IKB} as described¹⁶. mE3RS^{IKB} was PCR-amplified from a mouse expressed sequence tag clone, W61438, into the Flag vector. The mouse protein is 98% identical to human β -TrCP.

In vitro binding and ubiquitination assays. To incorporate ³⁵S-radiolabelled IKB α in the cellular NF- κ B complex, 6 μ l of an *in vitro*-translated, haemagglutinin-tagged IKB α (ref. 14) were incubated with 1 mg HeLa lysate prepared as described previously¹³. After 90 min incubation at 30 °C, IKB α /NF- κ B complexes were immunopurified as follows: 1 μ l of anti-p65 immunobeads was added, and the slurry was agitated for 2 h at 4 °C. The immune complexes were washed extensively with buffer A (1 M KCl, 0.5% NP40, 50 mM Tris buffer, pH 7.6, 1 mM DTT) and buffer B (50 mM Tris buffer, pH 7.6, 1 mM DTT) and incubated with constitutively active recombinant IKK- β , IKK2-EE (ref. 6), in the presence or absence of 2 mM ATP for 30 min at 37 °C; the immunobeads were eluted off with 1 mg ml⁻¹ p65 peptide. Mock phosphorylation was done by omitting the kinase or ATP. For conjugation assays the substrate was prepared from HeLa cells as described above. The E3 was prepared by the following procedure: 100 μ g of HeLa lysate supplemented with 1 μ M okadaic acid was mixed for 25 min at room temperature with immunobead-immobilized pIKB α , washed as above and eluted with the p65 peptide. Eluate fractions were supplemented with 0.2 μ g purified E1 and 1 μ g purified recombinant UBC5C incubated and analysed as described¹⁴.

Analysis of IKB-ubiquitin ligase fractions from stimulated HeLa cells was performed similarly, except that the cells were pre-incubated with the proteasome inhibitor ALLN (150 μ M) for 1 h and stimulated for 10 min with TNF. Slow-migrating IKB α -ubiquitin conjugates were detected by western blot using IKB α -specific antibodies. Ubiquitination assays with recombinant E3 proteins were done as follows: E3RS proteins were immunoprecipitated from transfected 293T cells with anti-Flag antibodies, eluted with 1 mg ml⁻¹ Flag peptide, combined with immunopurified, IKK- or mock-phosphorylated IKB α /NF- κ B complexes and assayed in a ubiquitination reaction as above. For the pIKB α binding assay, E3RS proteins from 293T transfected cells were immobilized on anti-Flag immunobeads and combined with the purified soluble IKB α /NF- κ B complexes. Bound proteins were washed extensively and identified by western blotting with anti-IKB α antibodies.

Immunopurification of E3RS^{IKB}. HeLa lysate (250 mg) was incubated with 250 μ l anti-p65 immunobeads. After four washes in buffer A (see above) and one wash in buffer B, half the beads were subject to *in vitro* phosphorylation with IKK and half underwent mock phosphorylation. The beads were washed twice in buffer B; 90% of each bead fraction was reincubated with half of the unbound HeLa lysate and 10% with 10 mg ³⁵S-metabolically labelled HeLa cell lysate (100 μ Ci ml⁻¹ Met/Cys for 8 h) for 30 min at 25 °C (with 1 μ M okadaic acid). The beads were then washed again four times with buffer A and once in buffer B and were eluted with 1 mg ml⁻¹ p65 peptide. Eluate fractions derived from both the hot and the cold lysates were mixed, boiled in SDS-sample buffer and analysed by 7.5% SDS-PAGE and autoradiography. Protein bands were reduced in-gel, S-alkylated and digested in-gel with an excess of trypsin as described¹⁵. Peptides were eluted with 1 μ l of 60% methanol, 5% formic acid directly into a nanoelectrospray needle¹⁵. Nanoelectrospray spectra were recorded on a quadrupole time-of-flight mass spectrometer (QqTOF, Perkin-Elmer Sciex, Toronto, Canada). Peptide-sequence tags were assembled from fragmentation spectra and searched against a non-redundant protein-sequence database (NRDB) maintained at the European Bioinformatics Institute (EBI, Hinxton Park, England) using the program PeptideSearch¹⁵.

Received 9 October; accepted 3 November 1998.

1. Baldwin, A. S. Jr The NF- κ B and IKB proteins: new discoveries and insights. *Annu. Rev. Immunol.* **14**, 649–683 (1996).
2. May, M. J. & Ghosh, S. Signal transduction through NF- κ B. *Immunol. Today* **19**, 80–88 (1998).
3. Baeuerle, P. A. & Baltimore, D. NF- κ B: ten years after. *Cell* **87**, 13–20 (1996).
4. DiDonato, J. A., Hayakawa, M., Rothwarf, D. M., Zandi, E. & Karin, M. A cytokine-responsive IKB kinase that activates the transcription factor NF- κ B. *Nature* **388**, 548–554 (1997).
5. Renner, C. H. *et al.* Identification and characterization of an IKB kinase. *Cell* **90**, 373–383 (1997).
6. Mercurio, F. *et al.* IKK-1 and IKK-2: cytokine-activated IKB kinases essential for NF- κ B activation. *Science* **278**, 860–866 (1997).

7. Woronicz, J. D., Gao, X., Cao, Z., Rothe, M. & Goeddel, D. V. IKB kinase- β : NF- κ B activation and complex formation with IKB kinase- α and NIK. *Science* **278**, 866–869 (1997).
8. Zandi, E., Rothwarf, D. M., Delhase, M., Hayakawa, M. & Karin, M. The IKB kinase complex (IKK) contains two kinase subunits, IKK α and IKK β , necessary for IKB phosphorylation and NF- κ B activation. *Cell* **91**, 243–252 (1997).
9. Yamaoka, S. *et al.* Complementation cloning of NEMO, a component of the IKB kinase complex essential for NF- κ B activation. *Cell* **93**, 1231–1240 (1998).
10. Rothwarf, D. M., Zandi, E., Natoli, G. & Karin, M. IKK- γ is an essential regulatory subunit of the IKB kinase complex. *Nature* **395**, 297–300 (1998).
11. Cohen, L., Henzel, W. J. & Baeuerle, P. A. IKAP is a scaffold protein of the IKB kinase complex. *Nature* **395**, 292–296 (1998).
12. Chen, Z. *et al.* Signal-induced site-specific phosphorylation targets IKB α to the ubiquitin-proteasome pathway. *Genes Dev.* **9**, 1586–1597 (1995).
13. Alkalay, I. *et al.* Stimulation-dependent IKB α phosphorylation marks the NF- κ B inhibitor for degradation via the ubiquitin-proteasome pathway. *Proc. Natl. Acad. Sci. USA* **92**, 10599–10603 (1995).
14. Yaron, A. *et al.* Inhibition of NF- κ B cellular function via specific targeting of the IKB-ubiquitin ligase. *EMBO J.* **16**, 6486–6494 (1997).
15. Wilm, M. *et al.* Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **379**, 466–469 (1996).
16. Margottin, F. *et al.* A novel human WD protein, h- β TrCP, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol. Cell* **1**, 565–574 (1998).
17. Shevchenko, A. *et al.* Rapid *de novo* peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer. *Rapid Commun. Mass Spectrom.* **11**, 1015–1024 (1997).
18. Herschko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479 (1998).
19. Hochstrasser, M. Ubiquitin-dependent protein degradation. *Annu. Rev. Genet.* **30**, 405–439 (1996).
20. Varshavsky, A. The ubiquitin system. *Trends Biochem. Sci.* **22**, 383–387 (1997).
21. Paul, M. & Jabbar, M. A. Phosphorylation of both phosphoacceptor sites in the HIV-1 Vpu cytoplasmic domain is essential for Vpu-mediated ER degradation of CD4. *Virology* **232**, 207–216 (1997).
22. Jiang, J. & Struhl, G. Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb. *Nature* **391**, 493–496 (1998).
23. Skowron, D., Craig, K. L., Tyers, M., Elledge, S. J. & Harper, J. W. F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell* **91**, 209–219 (1997).
24. Feldman, R. M., Correll, C. C., Kaplan, K. B. & Deshaies, R. J. A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell* **91**, 221–230 (1997).
25. Patton, E. E. *et al.* Cdc53 is a scaffold protein for multiple Cdc34/Skp1/F-box protein complexes that regulate cell division and methionine biosynthesis in yeast. *Genes Dev.* **12**, 692–705 (1998).
26. Elledge, S. J. & Harper, J. W. The role of protein stability in the cell cycle and cancer. *Biochim. Biophys. Acta* **1377**, M61–M70 (1998).
27. Jensen, J. P., Bates, P. W., Yang, M., Vierstra, R. D. & Weissman, A. M. Identification of a family of closely related human ubiquitin conjugating enzymes. *J. Biol. Chem.* **270**, 30408–30414 (1995).

Acknowledgements. We thank I. Alkalay for her invaluable role in developing many of the assays used in this work, A. Ciechanover for his support and advice in the early phases of the project, and A. Mahler and A. Bar-Sinai for comments on the manuscript. This research was supported by grants from the Israel Science Foundation funded by the Israel Academy for Sciences and Humanities-Centers of Excellence Program, the German-Israeli Foundation for Scientific Research and Development, and Signal Pharmaceuticals Inc. (San Diego).

Correspondence and requests for materials should be addressed to Y.B.-N. (ybinon@cc.huji.ac.il). The cDNA sequences for hE3RS^{IKB} and mE3RS^{IKB} have been deposited in Genbank under accession numbers AF101784 and AF099932, respectively.

Regulation of activity of the transcription factor GATA-1 by acetylation

Joan Boyes^{*†}, Peter Byfield[†], Yoshihiro Nakatani[‡] & Vasily Ogrzyko[‡]

^{*} Section of Gene Function and Regulation, Chester Beatty Laboratories at The Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK
[†] MRC Clinical Sciences Center, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK

[‡] Laboratory of Molecular Growth Regulation, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892-2753, USA

Modification of histones, DNA-binding proteins found in chromatin, by addition of acetyl groups occurs to a greater degree when the histones are associated with transcriptionally active DNA^{1,2}. A breakthrough in understanding how this acetylation is mediated was the discovery that various transcriptional co-activator proteins have intrinsic histone acetyltransferase activity (for example, Gcn5p (ref. 3), PCAF⁴, TAF_{II}250 (ref. 5) and p300/CBP^{6,7}). These acetyltransferases also modify certain transcription factors (TFIIIE β , TFIIIF, EKLF and p53 (refs 8–10)). GATA-1 is an important transcription factor in the haematopoietic lineage¹¹

and is essential for terminal differentiation of erythrocytes and megakaryocytes^{12,13}. It is associated *in vivo* with the acetyltransferase p300/CBP¹⁴. Here we report that GATA-1 is acetylated *in vitro* by p300. This significantly increases the amount of GATA-1 bound to DNA and alters the mobility of GATA-1-DNA complexes, suggestive of a conformational change in GATA-1. GATA-1 is also acetylated *in vivo* and acetylation directly stimulates GATA-1-dependent transcription. Mutagenesis of important acetylated residues shows that there is a relationship between the acetylation and *in vivo* function of GATA-1. We propose that acetylation of transcription factors can alter interactions between these factors

and DNA and among different transcription factors, and is an integral part of transcription and differentiation processes.

To determine whether GATA-1 is acetylated by a histone acetyltransferase, we incubated full-length chicken GATA-1 (cGATA-1) with highly purified p300 and [¹⁴C]acetyl CoA. cGATA-1 was significantly labelled whereas bovine serum albumin (BSA) remained unmodified (Fig. 1a). To verify that the acetylation was by p300 and not by proteins co-fractionating with cGATA-1, we incubated a high-performance liquid chromatography (HPLC)-purified peptide containing only the DNA-binding domain of GATA-1 (double-zinc-finger peptide¹⁵) with p300. This peptide

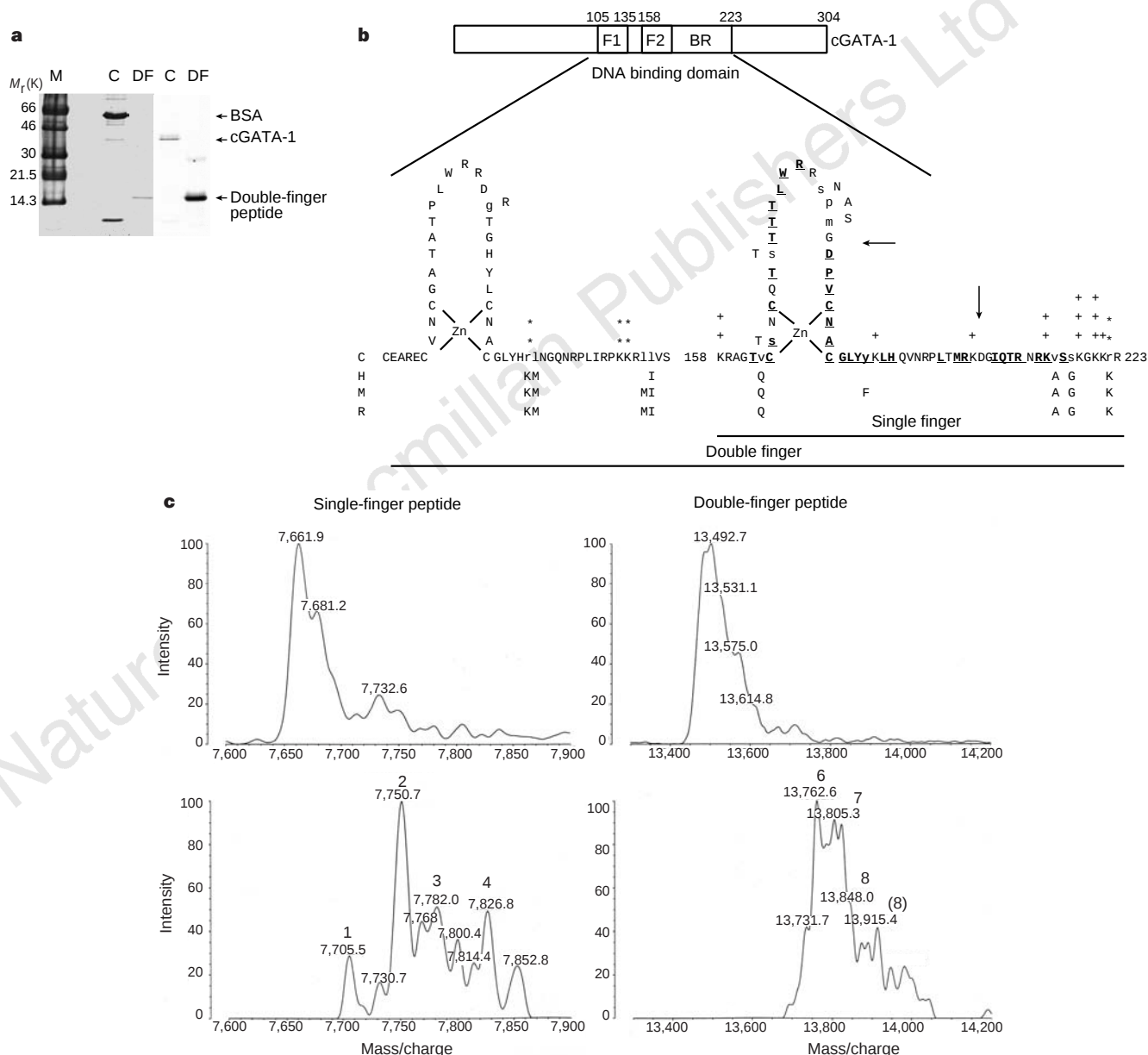


Figure 1 The DNA-binding domain of GATA-1 is acetylated by p300. **a**, Fractionated cGATA-1 (C) or double-finger peptide¹⁵ (DF) was incubated with p300 and [¹⁴C]acetyl CoA. The Coomassie-blue-stained gel is to the left; the autoradiograph is to the right. Lane M shows relative molecular mass standards. **b**, Amino acids spanning the chicken (C) GATA-1 DNA-binding domain are shown (F1 and F2, N- and C-terminal zinc fingers, respectively; BR, basic region). Changes in human (H), mouse (M) and rat (R) proteins are indicated. The DF peptide¹⁵ is from human GATA-1. Bold letters indicate residues shown to contact DNA²⁷; arrows indicate the points of cleavage by Asp-N; plus

symbols indicate relative levels of acetylation (+++ represents the highest level of acetylation); exact acetylation levels of lysines unique to the DF peptide were not determined but are relatively high (labelled with two asterisks arranged vertically). **c**, Mass spectra of the double-finger (right) and single-finger (left) peptides for their non-acetylated (upper panel) and acetylated (lower panel) forms. Numbers indicate [¹⁴C]acetyl groups incorporated. The *M_r* value for a [¹⁴C]acetyl group is 44. There was heterogeneity in the starting peptide and parentheses indicate incorporation of [¹⁴C]acetyl groups into a slightly larger peptide.

contains all the lysine residues, potential sites of acetylation, present in cGATA-1 (ref. 16) (Fig. 1b). Using these highly purified reagents, ^{14}C was efficiently incorporated (Fig. 1a).

To ascertain the extent of acetylation, we subjected non-acetylated and acetylated forms of the double-finger peptide to mass spectroscopy. We detected increases in mass consistent with the presence of at least seven markedly acetylated sites and one partially acetylated site (Fig. 1c). The positions of these acetylated lysines

were mapped using a smaller peptide covering the carboxy-terminal zinc finger and adjacent basic region (single-finger peptide¹⁷; Fig. 1b). This peptide contains seven of the eleven lysines of the double-finger peptide. Only two heavily and two partially acetylated sites were detected, indicating that all of the lysines present in the double-finger peptide that are not also in the single-finger peptide were probably at least partially acetylated (Fig. 1c).

We further studied the acetylated residues within the single-finger peptide by cleaving the acetylated peptide with the endoprotease Asp-N (Fig. 1b). This showed that the C-terminal fragment (containing four lysines) was the most acetylated fragment; there was also some labelling of the amino-terminal fragment (containing one lysine; see Supplementary Information). Radioactive sequencing pinpointed the positions of the acetylated sites within the C-terminal fragment: highest incorporation of label was into lysines 218 and 220, whereas lysine 214 was labelled to a lesser extent (see Supplementary Information). The positions of acetylated residues in both zinc fingers are shown in Fig. 1b. The sequences surrounding most of the sites of acetylation conform well to the consensus for acetylation of non-histone proteins (K/RXKK) suggested from studies of p53 (ref. 9). Mutagenesis studies (see below) confirmed a role for at least some of these acetylated lysines *in vivo*. We cannot exclude, however, the possibility that other acetylated sites are functional *in vivo* through the action of different acetyltransferases.

As all acetylated lysines were within the DNA-binding region of GATA-1, we next investigated the effects of acetylation on DNA binding. We studied the binding of acetylated and non-acetylated forms of the full-length protein and double-finger peptide to an oligonucleotide containing a single GATA-1-binding site. Acetylation of lysines would be expected to negatively affect DNA binding. Surprisingly, acetylation caused a distinct increase in GATA-1–DNA binding (Fig. 2a, b). Moreover, we observed two GATA-1–DNA complexes when non-acetylated GATA-1 was used, a faster-migrating complex and a more slowly migrating complex. Following acetylation, only a more slowly migrating complex was detected. This indicated that acetylation may have induced a conformational change in GATA-1. This idea was supported by a comparison of the mobility of the acetylated and non-acetylated slowly migrating forms: the acetylated form migrated slightly more quickly than the non-acetylated form. This was particularly apparent upon binding of the double-finger peptide to a palindromic GATA-binding site¹⁵, to which both zinc fingers bind^{15,18} (Fig. 2c), although the effect was also seen at single GATA-1-binding sites (to which only one zinc finger binds^{15,18,19}). The altered mobility at both types of binding site indicated that the DNA-binding conformations before and after acetylation are distinct. Moreover, as the DNA-binding domain is an important point of interaction with transcription factors such as Sp1 (ref. 20), the adapter protein FOG-1 (ref. 21) and the acetyltransferase CBP¹⁴, it is likely that if acetylation were indeed to induce a conformational change, it will alter protein–protein interactions as well as stimulating GATA-1–DNA interactions.

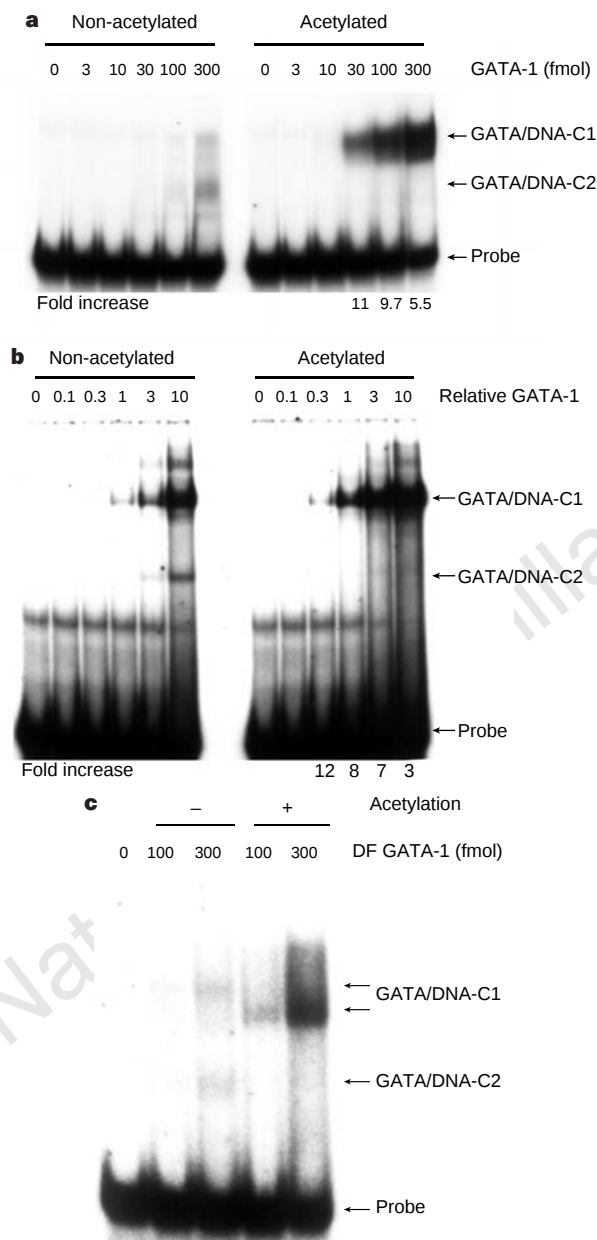


Figure 2 Binding of acetylated and non-acetylated GATA-1 to DNA. **a**, Acetylated and non-acetylated double-finger peptide generated two complexes (C1, the more slowly migrating complex, and C2, the more quickly migrating complex) with DNA carrying a simple WGATAR site. The fold increases of binding due to acetylation were calculated by comparing the increased percentage of probe bound at each concentration of acetylated relative to non-acetylated protein. **b**, Band-shift analysis of full-length acetylated or mock-acetylated cGATA-1, using the relative volumes of GATA-1 indicated. **c**, Acetylated and non-acetylated double-finger GATA-1 were bound to a consensus palindromic GATA-1-binding site¹⁵. Binding of both zinc fingers to this palindromic site generates a complex different to that produced at WGATAR sites^{15,18}. Acetylation results in a more quickly migrating complex, indicating that the acetylation-induced DNA-binding conformation is distinct from the conformation of non-acetylated GATA-1.

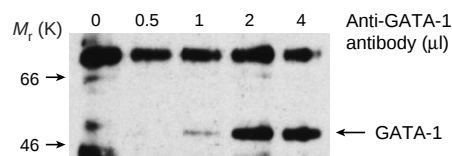


Figure 3 GATA-1 is acetylated *in vivo*. GATA-1 was immunoprecipitated from MEL cells using anti-mGATA-1 antibody. Acetylated proteins were detected using anti-acetyl-lysine antisera². Control experiments confirmed the species of M_r 50K as GATA-1 by probing with anti-GATA-1 antibodies. The anti-acetyl-lysine and anti-GATA-1 antibodies detected bands of similar intensities, indicating that a significant fraction of GATA-1 may be acetylated *in vivo*. The differing relative affinities of the antibodies, however, preclude definitive determination of acetylation levels.

To ascertain whether our *in vitro* observations are relevant *in vivo*, we next investigated whether GATA-1 is acetylated *in vivo*. We used antibodies against mouse GATA-1 (mGATA-1) to immunoprecipitate GATA-1 from mouse erythroleukaemia cells and determined the acetylation status of the immunoprecipitated protein by using anti-acetyl-lysine antisera². We detected mGATA-1 as a protein of relative molecular mass 50,000 (M_r 50K) (Fig. 3). We confirmed that this species was GATA-1 by reprobing western blots with a second anti-GATA-1 antibody. The same species was detected, showing that GATA-1 is indeed acetylated *in vivo* (data not shown).

To determine the *in vivo* functional consequences of GATA-1 acetylation, we studied the effects of acetylation on GATA-dependent transcription. NIH3T3 cells were transfected with a cGATA-1 expression plasmid and a reporter plasmid carrying a GATA-1-dependent promoter. GATA-dependent transcription was notably enhanced by co-expression of p300 (Fig. 4a). This is a direct effect of acetylation, as almost no transcriptional enhancement was seen when a p300 protein with a mutant acetyltransferase domain was expressed. The stimulation of transcription is specifically due to acetylation of GATA-1, as p300-dependent activation was reduced by mutation of the acetylated lysines in GATA-1. A correlation

between the level of reduction of GATA-1 acetylation and the reduction of activation by p300 was indicated by mutation of one partially acetylated lysine; this caused only a small decrease in p300-dependent activation (Fig. 4b). One mechanism by which acetylation might stimulate transcription is by increasing DNA binding by GATA-1; however, it is possible that acetylation might instead increase transactivation of GATA-1 by other factors. Control experiments showed that the mutant and wild-type proteins were expressed at similar levels and that DNA binding by the mutants *in vitro* was very similar to DNA binding by wild-type proteins (Fig. 4c, d).

Comparison of our data to the results of published studies of mutation of GATA-1 further supported the possibility that acetylation of GATA-1 has an *in vivo* function. The amino acids in region 218–221 of cGATA-1 are well conserved (Fig. 1b) and, we show here, encompass the highly acetylated amino acids (Fig. 1b). Removal of only these four amino acids in mGATA-1 reduced both transactivation and megakaryocyte differentiation²². These amino acids previously had no known function. They are not required for DNA binding^{17,19}, nuclear localization²² or direct interactions of GATA-1 with other transcription factors^{20,21}. Neverthe-

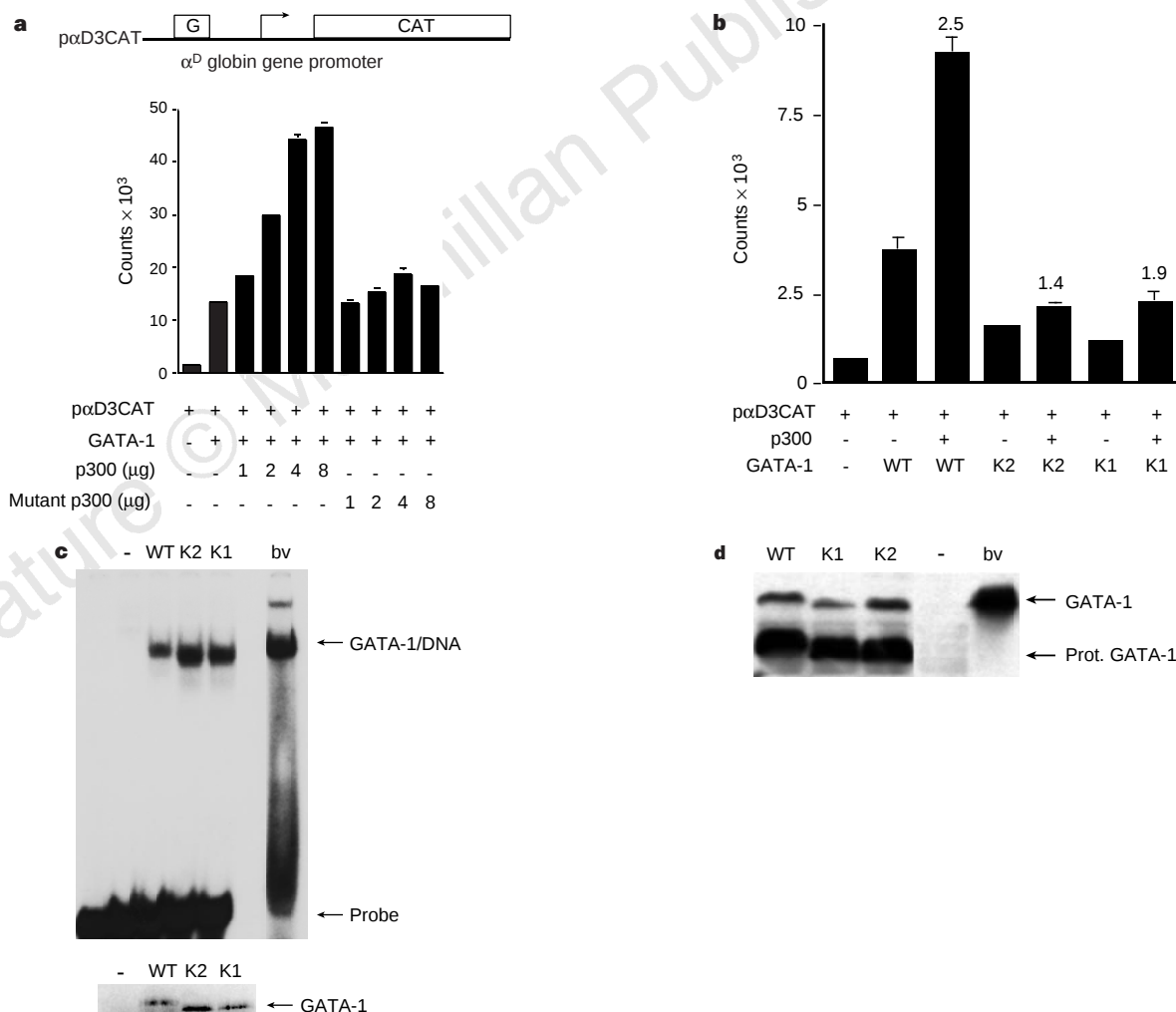


Figure 4 Acetylation of GATA-1 stimulated transcription directly *in vivo*. **a**, A CAT reporter construction, α D3CAT²⁶, carrying a single GATA-1-binding site (G) was co-transfected with a GATA-1 expression vector and expression vectors for p300 (wild-type or a mutant lacking the acetyltransferase domain). Control experiments confirmed that GATA-1 interacted equally with both p300 and the acetyltransferase deletion mutant. **b**, The acetylated lysines in cGATA-1 were mutated to arginines (K1 contains a mutation of K158; K2 contains a mutation of K151 and K152). The mutant GATA-1 proteins were used in the transfection assay described

in **a**. The fold stimulation by p300 is indicated by the numbers on the graph. **c**, Top, band-shift analysis of wild-type and mutant GATA-1 proteins prepared by *in vitro* translation; bv is purified GATA-1 from insect cells. Bottom, the relative amounts of protein generated. K1 and K2 migrated slightly differently from the wild-type protein; we suggest this is because of subtle alterations in folding, which do not impair DNA binding. **d**, Western blot analysis of proteins prepared from transfected cells. cGATA-1 is rapidly proteolysed during isolation from transfected cells and mainly yields proteolysed (prot.) GATA-1.

less, early deletion studies showed that these residues are important for haematopoietic differentiation²² and we have now shown that they are a major site of acetylation (Fig. 1b). These data therefore indicate a positive role for GATA-1 acetylation in haematopoietic differentiation. Future studies should verify the importance of GATA-1 acetylation in this process.

Our results show that acetylation directly regulates the transcriptional-activator function of GATA-1 *in vivo*. The data fulfill the prediction that non-histone proteins in addition to p53 are regulated by acetylation⁹. Our results also reveal that acetylation probably induces a conformational change in a key protein-protein interaction domain^{20,21}. This raises the possibility that a further function of acetyltransferases is to remodel protein-protein interactions. Modification of transcription factors by acetyltransferases could therefore be crucial in determining the architecture of promoter-bound transcription complexes and is yet another mechanism by which acetyltransferases can influence transcription and differentiation processes. □

Methods

Acetylation and DNA-binding reactions. Full-length p300 was expressed as a Flag-tagged fusion protein using a baculovirus transfer vector and was affinity-purified as described⁴. Purification of cGATA-1 from a baculoviral expression system²³ and the double-finger¹⁵ and single-finger¹⁷ have been described. Incorporation of [¹⁴C]acetyl groups was as follows: 20 µl cGATA-1 or 0.4 µg (30 pmol) double-finger peptide was incubated at 30 °C for 1 h with 5 nmol [¹⁴C]acetyl CoA (51 mCi mmol⁻¹, Amersham) and 120 ng p300 in acetyltransferase buffer⁶ in a final volume of 30 µl.

Acetylation with non-radioactive acetyl CoA for DNA-binding reactions was as follows: 1.5 pmol double-finger peptide or 1 µl cGATA-1 was incubated with 10 nmol acetyl CoA and 60 µg p300 as above. Acetyl CoA was omitted from mock acetylation reactions. Binding of double-finger peptide and cGATA-1 to DNA was studied by adding protein to 0.05 pmol probe (0.01 pmol for cGATA-1) in binding buffer¹⁵ and 0.1 µg poly dIdC. Electrophoresis was as described¹⁵. The oligonucleotide probes containing a single GATA-1-binding site and palindromic GATA-1-binding sites¹⁵ were 5'-AGCTTCGGTTCAGATAAA-CATTGAATTCA-3' and 5'-TCGCTATCAGATAAGGCCTTAT-3', respectively.

Mass spectroscopy and peptide sequencing. We labelled 80 µg single-finger peptide (10 nmol) using 30 nmol [¹⁴C]acetyl CoA and 0.9 µg p300 in acetylase buffer lacking phenylmethylsulphonylfluoride (PMSF) and dithiothreitol (DTT) for 140 min at 30 °C. The labelled peptide was cleaved with 0.4 µg Asp-N for 500 min at 37 °C. Asp-N activity was stopped by addition of EDTA and trifluoroacetic acid (TFA) to final concentrations of 5 mM and 3%, respectively. The proteolytic products were separated by reverse-phase HPLC on a C8 column using a gradient of acetonitrile in 0.1% (v/v) TFA. N-terminal Edman degradation was done using an applied Biosystems Model 473A machine. Radioactive sequencing was achieved by collection and counting of the phenylthiohydantoin derivatives. Laser desorption/ionization time of flight mass spectroscopy was done using a Compact Maldi 4 machine (Kratos Analytical).

Immunoprecipitations and western blotting. Unless indicated, all procedures were at 4 °C. We washed 5 × 10⁷ mouse erythroleukaemia cells in PBS and immunoprecipitated mGATA-1 as described²⁴ using an anti-GATA-1 antibody (N6, Santa Cruz Biotechnology). Immunoprecipitated material was electrophoresed on a 12% Laemmli gel. Western blotting, onto a poly(vinylidene fluoride) (PDVF) membrane, was at 0.67 mA cm⁻² for 1 h in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol). The membrane was blocked with 15% dried milk powder in PBS/0.1% Tween for 30 min and incubated with polyclonal rabbit anti-acetyl-lysine antisera² in PBS/0.1% Tween/3% BSA for 1 h. Bound antibodies were detected with peroxidase-conjugated goat anti-rabbit antibody (DAKO) and enhanced chemiluminescent substrate (Pierce). Antibodies were stripped using 2% SDS, 62.5 mM Tris, pH 6.8, 100 mM β-mercaptoethanol at 60 °C for 30 min. The membrane was reprobed with polyclonal goat anti-GATA antibody (M-20; Santa Cruz Biotechnology) for 1 h at 25 °C. Detection was with horse anti-goat antibody (Vector) and enhanced chemiluminescent substrate. Western blotting of transfected GATA-1 protein was done as described²⁵.

Transfections. The GATA-1 expression and reporter constructs, pRSV20-2 (ref. 16) and pαD3CAT²⁶, have been described. The vectors expressing mutant GATA-1 proteins (K1, K2) were prepared using overlapping polymerase chain reaction (PCR) primers with single base changes to convert the indicated lysines to arginines. The expression vectors containing full-length p300 (pCl.p300) and its acetyltransferase deletion derivative (pCl.p300HAT) were constructed by inserting a *XhoI*-*NotI* fragment carrying the N-terminal Flag-tagged p300 open reading frame or its acetyltransferase deletion derivative (Δ1,472-1,522) into pCl vector (Promega). NIH3T3 cells were grown in DMEM supplemented with 10% FCS. Cells were transfected using lipofectamine (Gibco; Fig. 4a) or calcium phosphate (Fig. 4b) with 10 µg DNA per 10-cm plate; total amounts of DNA were normalized using either pCl vector (to substitute for pRSV20-2) or a derivative of pCl.p300 with a frameshift at the start of the open reading frame (to substitute for pCl.p300 or pCl.p300HAT). Cells were collected 24-48 h after transfection and chloramphenicol acetyltransferase (CAT) assays were performed using the Promega CAT assay system.

Received 12 August; accepted 27 October 1998.

1. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349-352 (1997).
2. Hebbes, T. R., Thorne, A. W. & Crane-Robinson, C. A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* **7**, 1395-1402 (1988).
3. Brownell, J. E. *et al.* Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843-851 (1996).
4. Yang, X.-J., Ogryzko, V. V., Nishikawa, J., Howard, B. & Nakatani, Y. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1a. *Nature* **382**, 319-324 (1996).
5. Mizzen, C. A. *et al.* The TAF_{II}250 subunit of TFIID has histone acetyltransferase activity. *Cell* **87**, 1261-1270 (1996).
6. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953-959 (1996).
7. Bannister, A. J. & Kouzarides, T. The CBP coactivator is a histone acetyltransferase. *Nature* **384**, 641-643 (1996).
8. Imhof, A. *et al.* Acetylation of general transcription factors by histone acetyltransferases. *Curr. Biol.* **7**, 689-692 (1997).
9. Gu, W. & Roeder, R. G. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* **90**, 595-606 (1997).
10. Zhang, W. & Bieker, J. J. Acetylation and modulation of erythroid Kruppel-like factor (EKLF) activity by interaction with histone acetyltransferases. *Proc. Natl Acad. Sci. USA* **95**, 9855-9860 (1998).
11. Orkin, S. H. GATA-binding transcription factors in hematopoietic cells. *Blood* **80**, 575-581 (1992).
12. Pevny, L. *et al.* Erythroid differentiation in chimeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349**, 257-260 (1991).
13. Shivdasani, R. A., Fujiwara, Y., McDevitt, M. A. & Orkin, S. H. A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *EMBO J.* **16**, 3965-3973 (1997).
14. Blobel, G. A., Nakajima, T., Eckner, R., Montminy, M. & Orkin, S. H. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc. Natl Acad. Sci. USA* **95**, 2061-2066 (1998).
15. Trainor, C. D. *et al.* A palindromic regulatory site within vertebrate GATA-1 promoters requires both zinc fingers of the GATA-1 DNA-binding domain for high-affinity interaction. *Mol. Cell. Biol.* **16**, 2238-2247 (1996).
16. Evans, T. & Felsenfeld, G. The erythroid-specific transcription factor Eryf1: a new finger protein. *Cell* **68**, 597-612 (1989).
17. Omichinski, J. G. *et al.* A small single-"finger" peptide from the erythroid transcription factor GATA-1 binds specifically to DNA as a zinc or iron complex. *Proc. Natl Acad. Sci. USA* **90**, 1676-1680 (1993).
18. Yang, H.-Y. & Evans, T. Distinct roles for the two GATA-1 finger domains. *Mol. Cell. Biol.* **12**, 4562-4570 (1992).
19. Martin, D. I. K. & Orkin, S. H. Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf1. *Genes Dev.* **4**, 1886-1898 (1990).
20. Merika, M. & Orkin, S. H. Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Kruppel family proteins Sp1 and EKLF. *Mol. Cell. Biol.* **15**, 2437-2447 (1995).
21. Tsang, A. P. *et al.* FOG, a multiple zinc finger protein acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocyte differentiation. *Cell* **90**, 109-119 (1997).
22. Visvader, J. E., Crossley, M., Hill, J., Orkin, S. H. & Adams, J. M. The C-terminal zinc finger of GATA-1 or GATA-2 is sufficient to induce megakaryocyte differentiation of an early myeloid cell line. *Mol. Cell. Biol.* **15**, 634-641 (1995).
23. Pikaart, M. J. & Felsenfeld, G. Expression and codon usage optimisation of the erythroid-specific transcription factor cGATA-1 in baculoviral and bacterial systems. *Protein Expr. Purif.* **8**, 469-475 (1996).
24. Towatari, M. *et al.* Regulation of GATA-2 phosphorylation by mitogen-activated protein kinase and interleukin-3. *J. Biol. Chem.* **270**, 4101-4107 (1995).
25. Briegel, K. *et al.* Regulation and function of transcription factor GATA-1 during red blood cell differentiation. *Development* **122**, 3839-3850 (1996).
26. Evans, T. & Felsenfeld, G. Trans-activation of a globin promoter in nonerythroid cells. *Mol. Cell. Biol.* **11**, 843-853 (1991).
27. Omichinski, J. G. *et al.* NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* **261**, 438-446 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank N. Dillon and M. Greaves, in whose laboratories some of this work was done, for support and comments; C. Crane-Robinson for the anti-acetyl-lysine antisera; J. Omichinski for the GATA-1 peptides; M. Zenke for anti-cGATA-1 antisera; M. Pikaart and M. Towatari for plasmids; M. Antoniou for MEL cells; and A. Ashworth, T. Enver, G. Felsenfeld, A. Fisher, J. Mermoud and E. Milot for discussions and comments on the manuscript. J.B. acknowledges support from the Kay Kendall Leukaemia Fund.

Correspondence and requests for materials should be addressed to J.B. (e-mail: jboyes@icr.ac.uk).

Cell biology smorgasbord

Finding the answers to complex life science problems requires implements well matched to their task— such tools may be found in the emerging electronic marketplaces that operate over the the Internet and kick off this cell biology section.

BioSupplyNet

From SciQuest

www.biosupplynet.com

An Internet marketplace that provides scientific supplies to life science researchers

While BioSupplyNet will continue to publish its annual print resource, the *Source Book*, and run its Internet-based product sourcing website, it is now a wholly-owned subsidiary of SciQuest. BioSupplyNet has a joint marketing and distribution agreement with Cold Spring Harbor Laboratory. SciQuest offers scientists from all market sectors a broad and deep collection of laboratory supplies through the Internet. Whereas the SciQuest website (www.sciquest.com) serves as a product search portal for pharmaceutical and industrial laboratory scientists, the BioSupplyNet website is structured to be a web-based product directory for the biomedical research community and to serve as a gateway for university, government and biotechnology researchers. BioSupplyNet was formed in 1996 to expand commercially the print and Internet product directories developed by Cold Spring Harbor Laboratory, and to complement its laboratory techniques manuals.

Reader Enquiry No. 100

Chemdex.com

From Chemdex

www.chemdex.com

Providing electronic commerce solutions for the life science research industry

Offering more than 100 suppliers of biological and chemical reagents, Chemdex.com offers products through its full-transactional website. Suppliers who have signed on with this service include Amersham Pharmacia Biotech, Chemicon, CN Biosciences, Molecular Probes, Stratagene and Zymed Laboratories (a complete list of suppliers is on the website). Soon, more than 200,000 products will be available for online purchase. Moreover, this purchasing system may help to eliminate the need to search through multiple catalogues, as well as placing numerous phone calls, to buy the products from the various suppliers that researchers already know and trust. Combining a very large product selection, a powerful search engine and a streamlined electronic ordering process, this virtual warehouse allows users to order multiple products quickly and easily from a range of vendors. Chemdex is available to individuals and to institutions, in the later case, providing purchasing solutions for pharmaceutical and

biotechnology corporations, and for academic or government institutions. These solutions integrate the functionality of the service into existing purchasing processes, offering customized options like requisition routing and approval, summary billing, consolidated reporting and the set-up and management of pre-negotiated discount pricing from suppliers. The system also provides a history of purchasing for users, simplifying the re-ordering process, as well as record keeping. Through automated work-flow approval processes, consolidated and automated invoicing and billing, and systems integration, these products bring cost reductions and more effective purchasing to Chemdex customers.

Reader Enquiry No. 101

Imaging

BioScope AFM

From Digital Instruments

www.di.com

An atomic-force microscope (AFM) designed for bioscience research using inverted optical microscopes

New capabilities include micropositioning adjustments, which have been added for fast, precise alignment of the AFM probe along the optical axis of the inverted optical microscope. There are also optional manual and motorized precision stages with 2.54-cm travel and submicrometre resolution available for precise positioning of the AFM probe and for scanning. An optical image-capture capability has been added, which allows the AFM and optical images to be viewed and saved for further processing and printing, or for exporting to presentations and publications. In addition, a new integrated acoustic/vibration



Steady probing: Digital Instruments' BioScope.

isolation system is available to reduce the signal-to-noise ratio, even in noisy laboratory environments. The rigid design of the BioScope now allows the examination of structures with nanometre-scale resolution and tip positioning capability. The recent development of a TappingMode, phase imaging and force-volume imaging for fluid operation further enhance resolution for imaging in biologically-relevant environments.

Reader Enquiry No. 102

MRC-1024ES

From Bio-Rad

www.bio-rad.com

Extended imaging capabilities that include four-dimensional imaging

The new MRC-1024ES includes LaserSharp software, featuring TimeCourse options to provide four-dimensional imaging: a sequence of three-dimensional data sets (Z series) is collected over an extended period of time to record such processes as mitosis, gene transcription, differentiation and morphogenesis. Additionally, new co-localization software generates a variety of output image masks, co-localization coefficients and a two-dimensional fluorogram. Low-cost printer options are now a standard part of the hardware features of the MRC-1024ES, as are the new high-resolution, z-drive and the GFP filter-block option. It is the company's policy to provide users with upgrade opportunities; thus, the MRC-1024ES is now also available as a multi-photon-ready system. This allows users to buy a confocal system with all the necessary optical components for multi-photon imaging already in place, thus allowing on-site upgrades to a full multi-photon imaging system when required.

Reader Enquiry No. 103

CK40-RFA

From Olympus

www.olympus.de

A new incident-light, fluorescence attachment to extend the applications of the CK40 microscope to include fluorescence

Fitting into the standard frame, users can either incorporate this unit with a newly purchased CK40 or economically upgrade their system at a later date to meet the changing requirements of their laboratory. The CK40-RFA fluorescence unit provides cost-effective fluorescence observation for educational, clinical or research work. Standard objectives can be used with the

unit, which may prove invaluable to users undertaking examination of specimens by the fluorescence antibody method or the auto-fluorescence of biological specimens. For special cases, particularly with $\times 40$ magnification, a long-working-distance CPlan achromat objective with a numerical aperture of 0.55 is offered. Two separate filter modules are available for green and blue excitation, as standard; these can be moved easily in and out of the light path to allow rapid changeover between blue or green excitation and transmitted-light observation. By mounting both the green and blue excitation filter units on the CK40, sequential observation of double stained specimens can be quickly obtained. Optionally, an empty cube can be offered for unusual filter combinations depending on customer requests. The fluorescence unit uses a 50 W mercury light source with a centring mechanism and a reflector mirror in the lamp housing to optimize brightness. The illuminator accommodates the use of neutral-density filters and has a built-in, field-iris-diaphragm centring mechanism with a shutter.

Reader Enquiry No. 104

UltraPix 2000

From Life Science Resources www.lsr.co.uk

A programmable, high-speed, colour and monochromatic digital imaging camera

This is a high-resolution, cooled digital camera utilizing progressive-scan, interline technology to provide enhanced spectral sensitivity in the blue region of the spectrum. UltraPix 2000 should prove useful in low-light-level fluorescence and bright-field applications where high-quality, static or dynamic images are required for biological and medical research. Optimized specifically for use on microscopes, the camera is compact, fast and flexible, offering both full-colour and black-and-white imaging at full-chip resolution. A user-selectable RGB filter for colour imaging is optionally available, making it well suited for use in busy, multidisciplinary laboratories that require both monochrome and full-colour imaging. In addition, resolution of $1,280 \times 1,024$ pixels at 12-bit depth offers a wide dynamic range. UltraPix 2000 is controlled by the UltraPlus camera interface, providing an advanced imaging environment. Automatic exposure and autoscale enable a quick and easy set-up. These features combined with sub-array selection and on-chip, asymmetric binning to make high-throughput, high-speed image acquisition possible. Eight or more images, each at a different wavelength, can be captured and then displayed in the thumbnail view gallery before image overlay to show co-localization of up to eight probes. Morphometric measurements and

densitometric, quantitative analysis can be applied to up to 100 user-defined regions of interest within a chosen image.

Reader Enquiry No. 105

Cell growth and disaggregation

BioFlo 4500 fermentor

From New Brunswick Scientific

www.nbsc.com

A fully customizable benchtop fermentor for microbial, insect and mammalian cell culture

In applications ranging from process development to small-scale production, the BioFlo 4500 fermentor offers advanced control capabilities through a built-in multi-loop controller with a touch-screen interface. The system's vessel can be sterilized without disassembly. Onscreen entry of new process control loops, sterilization cycles and calibration routines is intuitive, making the BioFlo 4500 simple to operate. Unique capabilities include rapid temperature shifts for heat induction, oxygen transfer rates of up to $350 \text{ mM O}_2 \text{ l}^{-1} \text{ h}^{-1}$ (even with highly viscous cultures) and, for scale-up studies, emulation of much larger bioreactors through programmable heating, cooling and sterilization valve sequencing.

Reader Enquiry No. 106

MediMachine

From BDIS

www.bdis.com

The company now offers Consul TS's Medimachine system for tissue disaggregation

According to Becton Dickinson Immunocytometry Systems, the unit has been designed for ease of use, as well as to be a rapid, safe means of obtaining single-cell suspensions or nuclei from tissue samples for further analysis by FACS brand flow cytometers, or other molecular biology approaches such as PCR. This system should increase the efficiency of research laboratories while simultaneously reducing the risk of accidental exposure associated with other methods currently employed for tissue disaggregation. Increasing numbers of research laboratories are now performing flow cytometric analysis of solid tissues and tumours for



The MediMachine is now offered by BDIS.

cellular antigen expression, DNA ploidy and proliferative index. This compact device uses Medicons and Filcons to disaggregate and filter samples. After disaggregation and filtration, the sample is ready for processing using routine research procedures, such as cell analysis or DNA isolation.

Reader Enquiry No. 107

Assays and cytometry

FlowTACS

From Genzyme Diagnostics

www.genzyme.com

A new apoptosis detection kit for flow cytometry

FlowTACS uses the company's proprietary labelling systems to maximize the signal of apoptotic cells versus that of necrotic cells, minimizing nonspecific background labelling. FlowTACS includes Genzyme's non-toxic and signal-enhancing cation buffering systems to optimize apoptosis detection. The CytoPore permeabilization reagent has been formulated to maintain the integrity of DNA and protein antigens and facilitates multi-colour analysis by immunohistochemistry. FlowTACS is a quantitative, rapid assay and contains all of the reagents necessary for start-to-finish sample labelling. Each kit also includes TACS-Nuclease to generate a sample-specific positive control by creating DNA breaks.

Reader Enquiry No. 108

TSA-enhanced fluo-peptides

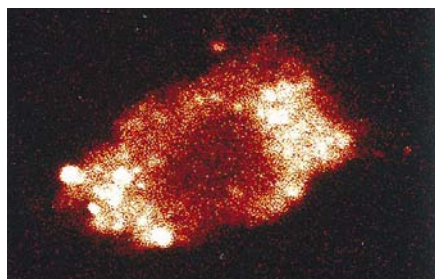
From NEN

www.nenlife.com

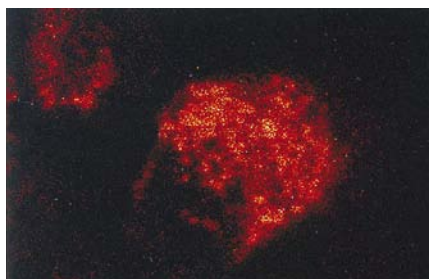
Formulated to boost the sensitivity of fluorescence detection in G-protein-coupled receptors (GPCRs)

The manufacturer now offers protocols describing the applications for TSA (tyramide signal amplification) with fluo-peptides. (These are biologically active, fluorescein-conjugated peptides from Advanced Bioconcept). Fluo-peptides are high-affinity fluorescent probes for the characterization of GPCRs, allowing the system to detect low-level, receptor-expression signals by making them become stronger and clearer. Protocols are available for both direct and indirect cellular imaging applications. Fluo-peptides were created to enable high-resolution visualization of receptors on cell cultures, frozen sections, 'live-slice' preparations and *in vivo* specimens. According to the manufacturer, TSA enhances the signal of all proteins and nucleic acids in immunohistochemistry (IHC) and *in situ* hybridization up to 100-fold. It can be used with either chromogenic or fluorescent detection systems, with no loss of resolution.

Reader Enquiry No. 109



TSA-enhanced fluo-peptides (left) make it possible to extend the detectable range of receptor-expression levels to facilitate the visualization in multiple spectra.



nflux

From Molecular Probes www.probes.com

A pinocytic cell-loading reagent

According to the manufacturer, users can now rapidly load membrane impermeant compounds into live cells without resorting to invasive procedures, such as microinjection or electroporation. This pinocytic cell-loading reagent helps to permeate cells with a diverse array of compounds, including DNA, nucleic acid stains, fluorescent ion indicators, biocytins, hydrazides, dextrans and cytoskeletal proteins. It can be used with cells grown on coverslips, in suspension or in tissue culture flasks. Moreover, the reagent employs a simple technique, based on the osmotic lysis of pinocytic vesicles, which have been used to load compounds into endothelial, carcinoma and leukemia cells, fibroblasts, macrophages and T cells. QuickTime movies of the reagent in action, can be seen on the company's website: <www.probes.com/lit/feature/influx>.

Reader Enquiry No. 110

Blasticidin

From Invitrogen www.invitrogen.com

An antibiotic for rapid generation of stable mammalian cell lines

This is a nucleoside antibiotic that inhibits protein synthesis in both prokaryotic and eukaryotic systems. It is said to be an extremely potent selection agent for cultured mammalian cells. Complete cell death is often achieved within 5–7 days of treatment at concentrations as low as 1–2.5 $\mu\text{g ml}^{-1}$. This is much faster than the three to four weeks required by other antibiotics to generate stable cell lines, says Invitrogen. The company's pcDNA6 series of EpiTag expression vectors confer resistance to blasticidin and allow users to take full advantage of this antibiotic.

Reader Enquiry No. 111

ITPMPA

From RBI www.callrbi.com

According to RBI, this is currently the only known selective GABA_C receptor antagonist

Licensed from the University of California

to RBI, TPMPA [(1,2,5,6-tetrahydropyridine-4-yl) methylphosphonic acid] is said to be the first highly selective ($K_b=2.1 \mu\text{M}$) GABA_C receptor antagonist, exhibiting little affinity for GABA_A ($K_b=320 \mu\text{M}$) or GABA_B ($\text{EC}_{50}=500 \mu\text{M}$) receptors. Although there is increasing evidence to indicate that GABA_C receptors are present in various brain regions, current information on GABA_C receptors comes mainly from studies performed in the visual system, particularly in the retina. TPMPA provides a useful tool for studying the function of GABA_C receptors and will aid in identifying GABA_C receptor-mediated responses in other parts of the CNS.

Reader Enquiry No. 112

Lightwave spectrophotometer

From WPA www.wpaltid.co.uk

A low-cost ultraviolet/visible spectrophotometer for molecular biology

This new diode-array spectrophotometer should prove useful in routine determinations of purity and concentration of DNA or RNA. All wavelengths of light are measured simultaneously, allowing instantaneous display of $A_{260/280}$ ratios. The wavelength range of the instrument covers 200–825 nm, ensuring that bacterial growth rates can be measured. With instantaneous scanning and no moving parts a robust light-weight construction and a unique space-saving profile,



WPA's cell and molecular spectrometer.

Lightwave is well suited to the confines of the modern laboratory. It uses menu-driven, microprocessor controls with a large, easy-to-read screen that provides a full display of measurement data scans in a variety of formats. Up to 99 user-programmable methods can be stored and an RS232 port allows data to be printed out or downloaded to a PC for further analysis. The instrument is easy to clean and is designed for a long maintenance-free life without the need for recalibration. According to the manufacturer, the sample compartment is quickly and easily accessed, and requires no cover as the instrument is unaffected by stray light. Cuvettes with pathlengths up to 40 mm and test tubes up to 16-mm in diameter can be loaded, as can ultra-micro-cuvettes down to 70 μl . An integral cuvette tray acts both as a convenient storage area and as a rack during analysis.

Reader Enquiry No. 113

Gamma-B 25-hydroxy vitamin D assay

From IDS www.idstid.com

Designed to optimize precision and sensitivity of the 25-hydroxy vitamin D assay

An initial sample extraction step is followed by a standard assay procedure, with results that are stated to be available in about three hours. This new radioimmunoassay measures both native vitamin D₃ and synthetic vitamin D₂. It is said to offer excellent accuracy and precision both within and between assay (<8% and <12%, respectively), and is co-specific for 25-hydroxy vitamin D₂ and D₃. The assay's sensitivity is <1.2 ng ml⁻¹ (<3 nmol l⁻¹) and offers a convenient working range of 1.6–160 ng ml⁻¹ (4–400 nmol l⁻¹), says IDS. Together with high throughput and a rapid turnaround time, IDS has made it possible to measure the biologically active metabolite of vitamin D accurately in a simple assay, which can be easily assimilated into routine laboratory testing. The kit is available in a 100-tube format to accommodate the needs of most clinical and research laboratories. Moreover, it

ADVERTISEMENT



FULLY AUTOMATED DNA SEQUENCER

www.beckmancoulter.com/ceq20

READER ENQUIRY NO. 47

complements the existing Gamma-B 1,25-dihydroxy vitamin D kit and should prove useful as an addition to the company's range of bone and mineral metabolism products.

Reader Enquiry No. 114

DeadEnd

From Promega UK www.euro.promega.com/uk

Colorimetric apoptosis detection in a non-radioactive system

Designed to provide simple, accurate and rapid *in situ* detection of apoptic cells either at the single-cell level or in tissue sections. Fragmented DNA of apoptic cells is end-labelled using a modified, TdT-mediated, dUTP nick-end labelling assay. Biotinylated nucleotide is incorporated at the 3'-hydroxyl ends using the terminal deoxynucleotide transferase (TdT) enzyme. Streptavidin-HRP is then bound to these biotinylated nucleotides, which are detected using the stable chromagen diaminobenzidine (DAB). With this procedure, apoptic nuclei are stained brown. According to Promega, the assay complements its range of apoptosis detection systems, which includes both the fluorometric CaspACE assay system and the apoptosis detection system, which are applicable to cancer research, developmental biology of nervous and immune systems and routine pathology.

Reader Enquiry No. 115

PIPs and IsoPPs

From Echelon Research Laboratories

www.echelon-inc.com

A selection of soluble, affinity-labelled phosphoinositide polyphosphates (PIP_ns, and isoprenoid diphosphates (IsoPPs)

Designed for signal transduction research, researchers can select PIP_ns with various diacylglycerol chains, such as short-chain water-soluble analogues, intermediate-chain amphiphatic analogues, long-chain micelle-forming analogues or functionalized 'tetherable' analogues. These chemically synthesized products are guaranteed to be >95% pure. Prepackaged reagent kits are available at a discount, and custom IsoPPs and PIP_n affinity probes are also available — a selection of fluorescent PIP_ns will be released in winter 1998.

Reader Enquiry No. 116

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

correction

The URL as it appeared in Reader Enquiry No. 100 of Nature (396, 287; 1998)

"www.taconnic.co.uk" should have read

"www.taconnic.com".

ADVERTISEMENTS

Homogenizers?

- Tissue-Tearor™
- BeadBeater™
- Mini-BeadBeater™
- BioHomogenizer™

Check Us Out

www.biospec.com

BioSpec Products, Inc.
Ph/Fax 918-336-3363
e-mail info@biospec.com

New Probe Sizes
0.45cm!

READER ENQUIRY NO. 54

Detect residual DNA with Wako's DNA Extractor Kit or Isolate Genomic DNA with DNA Extractor WB Kit

Both use the Nal Method
No Phenol No Chloroform

Wako

Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 39

Check out our homepage at <http://www.resgen.com>

Peptides
\$18/residue

Antibodies
monoclonal - \$8,500
polyclonal - \$1,195

Complete services for your custom peptide and polyclonal or monoclonal antibody needs.

Research Genetics, Inc.
U.S. or Canada 800-533-4363
U.K. 0-800-89-1393
FAX 256-536-9016

READER ENQUIRY NO. 37

ANTIBODIES
For Research & Routine Histology

Inflammation – Angiogenesis – Tumors
Leukocytes – Endothelial Cells
Pathological Markers

Custom Monoclonals and Polyclonals

BMA Biomedicals Ltd.
(800) 634-3183 (US)
Intl: +41 61 811 62 22
+41 61 811 60 06 (fax)
info@bma.ch

READER ENQUIRY NO. 28

NEW ANTIBODIES FOR UBIQUITIN AND PROTEASOME RESEARCH FROM AFFINITI

20S Proteasome α - & β -subunits: HC2, HC3, HC8, HC9, ι , XAPC7, ζ , HC10, HN3, Z, MECL1, LMP2, LMP7; 19S regulator ATPase & non-ATPase subunits; 11S regulator subunits: PA28 α , β and γ ; **Signalosome** subunits; PGP9.5; E2; Ub-conjugates

For data sheets or more information on AFFINITI's extensive range of products for ubiquitin and proteasome research visit our website at:

[HTTP://WWW.PROTEASOME.COM](http://www.proteasome.com)

AFFINITI Research Products Ltd., Mamhead Castle, EXETER, EX6 8HD, UK
Tel. +44/0 1626 891010; Fax: +44/0 1626 891090; URL: <http://www.affiniti-res.com>

READER ENQUIRY NO. 38

Severn Biotech Ltd

Customer Services

Comprehensive Custom Service for DNA, Peptides and Polyclonal antibodies from SBL

DNA Synthesis

- Oligo's 0.05 μ M-10 μ M
- Labelled Fluorescein, Biotin 6-FAM, Hex, Tet et al
- Purified HPLC/Cis

Peptide Synthesis

- Peptide Synthesis from 1-1000mgs
- HPLC & Mass Spec Verification
- HPLC Purification to 95% +
- Conjugation to Carrier Proteins
- Custom Polyclonal Antisera

For a Fast Efficient Service Reply Today

Severn Biotech Ltd, Unit 2, Park Lane, Stourport Road, Kidderminster, Worcs DY11 6TJ, UK
Tel: (01562) 825286 Fax: (01562) 825284
E-mail: 101626.2102@compuserve.com

READER ENQUIRY NO. 55
Nature © Macmillan Publishers Ltd 1998